

STIC-Biotech/ChemLib

From:

Sent: To:

Schnizer, Richard Wednesday, August 27, 2003 2:28 PM STIC-Biotech/ChemLib 09/965,594

Subject:

Please search the commercial and published application databases for polypeptide SEQ ID NOS: 1, 12, 14, 16, 18, 20, 22, and 26 from 09/965,594. Please also search for nucleic acids that could encode these polypeptides.

Thank you-

Richard Schnizer, Ph.D. Patent Examiner Art Unit 1635 CM1 12E17 703-306-5441 Mail Box CM1 11E12

Searcher:
Phone:
Location:
Date Picked Up:
Date Completed:
Searcher Prep/Review:
Clerical:
Online time:

TYPE OF SEARCH:
NA Sequences:
AA Sequences:
Structures:
Bibliographic:
Litigation:
Full text:
Patent Family:

vendor/cost (wi	nere applic.)
STN:	
DIALOG:	
Questel/Orbit:	
DRLink:	
Lexis/Nexis:	
Sequence Sys.:	
www/Internet:	
Other (specify):	

```
AR118696 Sequence 109331 Sequence 109331 Sequence 109331 Sequence AR118728 Sequence AR11872 Sequence BD091382 HCV culti 108294 Sequence 108294 Sequence 108294 Sequence 108294 Sequence 108294 Sequence 108294 Sequence 10823 Hepatitis C 11894 Sequence 9 11894 Sequence 9 181885 Sequenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF369218 Hepatitis
AF369215 Hepatitis
AF18686 Sequence
106434 Sequence
109328 Sequence 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AF369224 Hepatitis
AF369237 Hepatitis
AF511948 Hepatitis
AF369230 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            508263 gRNA of Hep
508264 cDNA of Hep
D10749 Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AR118692 Sequence
106440 Sequence 54
I09329 Sequence 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AX441176 Sequence
AX467113 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AX395309 Sequence
AX454818 Sequence
                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                             Description
                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HPCPLYPRE
AF387805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AF369218
AF369235
AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR118696
I09331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR118722
AR118723
BD091382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR301300
AR176483
BD080334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AF387806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AF369224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR166930
                                                                                                                                                                                                            em_htgo_hum:*
em_htgo_mus:*
em_htgo_other:*
                                                               em_htg_other:*
                        em_htg_hum:*
                                                                                        htg_mus:*
                                                                                                       em_htg_pln:*
em_htg_rod:*
                                                                                                                                                   em_htg_mam:*
em_htg_vrt:*
                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                                                                                                                                                                                                                                   %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5360
5360
5360
6785
6785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7310
7310
7310
8987
9185
9185
9185
9185
9379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9401
9609
9609
9618
9646
9693
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9401
                                                                                                                                                                                             em_sy:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    \\
2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

2000

                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULI 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    August 30, 2003, 19:18:33; Search time 2365.6 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                                                                                                                                                                                                    953
1 MAPITAYAQOTRGLLGCIIT......GVAKAVDFIPVESLETIMRS
                 GenCore version 5.1.6
Copyright (c) 1993 · 2003 Compugen Ltd.
                                                                                                   - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                        2888711 segs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                     0.5
                                                                                                                                                                                                                                                                                                                            Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Minimum DB scg length: 0 Maximum DB scg length: 20000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GenEmbl:*

1: 9b_ba:*

2: 9b_htg:*

3: 9b_in:*

9b_ov:*

9b_pat:*

9b_pl:*

9b_pl:*

9b_pl:*
                                                                                                                                                                                                                                  US-09-965-594-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9b_st: *
9b_sy: *
9b_un: *
9b_vi: *
em_ba: *
em_fun: *
em_fun: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          em_in:*
em_mu:*
em_om:*
em_ov:*
em_pat:*
em_ph:*
em_ph:*
em_ph:*
em_pl:*
em_cn:*
                                                                                                                                                                                                                                                                                                                 BLOSUM62
                                                                                                                                                                                                                                  Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                       OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database :
                                                                                                                                                                                                                                                                            Sednence:
                                                                                                                                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                                                                   ..
0
```

~

160 480 540

```
/ transl_table=11
/transl_table=11
/transl_table=11
/protein_id="cab38232.1"
/protein_id="cab38232.1"
/protein_id="cab38232.1"
/protein_id="cab38232.1"
/protein_id="cab38232.1"
/db_xrcf="cal=21714048"
/db_xrcf="cal=21714048"
/db_xrcf="cal=21714048"
/db_xrcf="cal=21714048"
/db_xrcf="cal=21714048"
/db_xrcf="cal=21714048"
/db_xrcf="cal=21714048"
/translation=12714048"
/translation=12714048"
/translation=12714048"
/translation=12714048"
/translation=12714048"
/translation=12714048"
/translation=12714048"
/translation=12714048"
/translation=12714148"
/translation=1271448"
/translation=12714148"
/translation=12714148"
/translation=12714148"
/translation=12714148"
/translation=12714148"
/translation=1271448"
/translation=1271448"
/translation=1271448"
/translation=1271448"
/translation=1271448"
/translation=1271448"
/trans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAT 06-JUL-2002
121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                                                                                     /note-"representative NS3/4a conformational antigen'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chien, D.Y., Arcangel, P., Tandeske, L., George-Nasciemento, C., Coit, D. and Medina-Selby, A. Immunoasasys for anti-hev antibodies Patent: Wo 01968PO.A. 1 20-DEC-2001; CHIRON CORPORATION (US)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2058
180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. >2058
/note-"unnamed protein product"
/codon_start-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2058 bp
Sequence 1 from Patent W00196870.
AX454818
AX454818.1 GI:21714047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-1 (1-182) x AX454818 (1-2058)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1. .2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.56e-68
946.00
100.00%
98.90%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            artificial sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   synthetic construct
synthetic construct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           633 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111111
541 AGGTCC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
   01
                                                            301
                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 2
AX454818
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                         a
                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                             òγ
                                                                                                                                                                                                                                                                                                                                                                                                                      G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                     ò
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRGVAKAVDETPVENLETTMRSPVFTUNSSPPVVPOSFQVAHHAPTGSGKSTKVPAA
YAAQGYKVLANPSVAATLGFGAYMSKAHGIDNIFFGVRITTGSPITYSTYGKELA
DGGCSGGAYDLIICDECHSTDAFSILGIGTVLDQAFTAGRILYVLATATPPGSVTVPH
PNIEVPALSTTGEIPEFYGRAIPLEVIKGGRHLIFCHISKKODELAAKLVALGINAVA
YRGLDVSVIPPIGDVVVVATDALMTGYTGDFDSVIDCNICVTGYVDFSLDPIFTIETI
TLPQDAVSRYOPRGRYGKGKPGIYREVAPGERPSSVLCECYDAGCAMWELIPA
ETTVRLRAYMYTPGLPVCQDHLEFWEGVFTGLIHIDAHFLSQTKGSGENLPYLVAYGN
YVCARRQAPPPSWDOMMKCLIRLKPTHGPTFLLYRLGAVQNRTTHTHYTKYIMTGM
SADLEVVTSTWVLYGGVLAALAAYCLSIGCVVIVGRVVLSGKPAIIPDREVLYREFDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /codon_start=1
//transl_table=11
/protein_id="cable=11
/protein_id="cable=11
/db_xref="G1:21066309"
/db_xref="G1:2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAT 18-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MetalaProlleThrAlaTyralaGlnGlnThrArgGlyLeuSeuGlyCyslie:leThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note="representative NS3/4a conformational antigen"
1. .>2058
                                                                                                                                                                                                                                                                                                Chien, D.Y., Arcangel, P., Tandeske, L., George-Nasciemento, C. Coit, D. and Mcdina-Selby, A.

For antigen/antibody combination assay
Patent: Wo 0196875-A 2 20-DEC-2001;
CHIRON CORPORATION (US)
                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2058
180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism-"synthetic construct"
/mol_type-"genomic DNA"
/db_xref-"taxon:32630"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note="unnamed protein product"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         425
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                      2058 bp
Sequence 2 from Patent W00196875.
AX395309
AX395309.1 GI:21066308
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-1 (1-182) x AX395309 (1-2058)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.56e-68
946.00
100.008
98.908
99.278
                                                                                                                                                                                 synthetic construct
synthetic construct
artificial sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         634 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Bost Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                         LOCUS
DEFINITION
ACCESTON
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ..
Q
                                                                                                                                                                                                                                                                          REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
   4X395309
                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
```

q

g

Ω

ò

q

ö

q

õ

ò

g

g

ò

ò

g

셤

ò g ò

ò

```
/codon_start=1
/product=*NS3 protease"
/product=*NS3 protease"
/protein_id=*AAK54543.1*
/db_xref=*C1:14150561*
/translation=*APITATAQOTHGLLGC1ITSLTGRDKNOVEGEVQIVSTAAQTFL
/translation=*APITATAQOTHGLLGCIITSLTGRDKNOVEGEVQIVSTAAQTFL
/translation=*APITATASPEGPYIQMTNNOKDLYGWPAPQGSRSLTPCTCGSS
DLYLVTRHADVIPWRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCT
RGVAKAVDFIPVENLETTHRS*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VRL 03-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            480
                                                                                                                                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                      62 ThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAspLys 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                 122 AspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       102 SerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgGly
                                                                                                                                                                                                                                                                                                                                                         GlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGly
                                                                                                                                                                                                                                                                                                              22 LeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AF369235 543 bp RNA linear VRL
Hepatitis C virus Pt.lY NS3 protease gene, partial cds.
AF369235
                                                                                                                                                           543
180
1
0
0
                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                106
                                                                                                                                                                                                                                          US-09-965-594-1 (1-182) x AF369218 (1-543)
<1. .>543
/note-"polyprotein"
                                                                                                               158 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AF369235.1 GI:14150594
                                                                                                                                                          1.89e-68
945.00
100.00%
99.45%
99.16%
                                                                                                                166 c
                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ser 182
                                                                                                                 ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     541 TCC
                                                                                                                                                  Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182
                                                                                                                                                                                                                                                                                                                                                          42
                                                                                                                                                                                                        Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                  Alignment :
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
AF369235
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
SOURCE
  CDS
                                                                                                                                                                         Score:
                                                                                                                                                                                                         Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΩD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 (bases 1 to 543)
Holland-Staley, C.A., Kovari, L.C., Golenberg, E.M., Pobursky, K.J. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               03-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                 360
                                                                                                                                                                                                                                                                                       140
                                                                                                                                                                                                                                                                                                            420
                                                                                                                                                                                                                                                                                                                                                                                 180
                                                                                                                                                                                                                                                                                                                                                                                                       540
                                                                                                                                                                      240
                                                                                                                                                                                             100
                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                          120
                                                                              120
                                                                                                                           180
                                                                                                    20
                                 9
                                                       SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValG]nIleValSerThrAla 40
                                                                                                                                                  90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Direct Staley, C.A., Kovari, L.C., Golenberg, E. and Mayers, D.I.. Direct Submission Submission Submission Infectious Disease Research, Henry Ford Health Systems, 2799 W. Grand Blvd. Rm 7045 E & R, Detroit, MI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genetic diversity and response to IFN of the NS3 protease gene clinical strains of the hepatitis C virus Arch. Virol. 147 (7), 1385-1406 (2002)
                                                                                                                                                                                                                                                                                    241 CAAGACCTTGTGGGCGGCCGGCTCGGAAGGTAGCCGATCATTGACACCTGCACTTGC
                                                                                                                                                                                                                                                      LysaspLeuValGlyTrpProAlaProGlnClySerArgSerLeuThrProCysThrCys
                                                                                                                                                                                                                                          GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg
                                                                    MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus Pt.4 NS3 protease gene, partial cds. AF369218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive strand viruses, Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism-"Hepatitis C v
/mol_type-"genomic RNA"
/isolate-"Pt.4"
/db_xref-"taxon:11103"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note="type: 1A"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                              ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AGGTCC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mayers, D.L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12111414
                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                             81
                                                                                                                                                                                                                                          101
                                                                                                                                                                                                                                                                                                                                                                                                                                                   541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 3
AF369218
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
```

οχ g ô Q

a

ογ

g

g

ö

```
DNA
                                                                                                                                                                                                                 Unclassified.
1 (bases 1 to 5360)
(Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 53 21-NOV-2000;
Location/Qualifiers
1. 5360
                                                                                                                                           5360 bp I
US 6150087.
                                                                                                                                                                                                                                                                                                  1145
                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-1 (1-182) x AR118686 (1-5360)
                                                                                                                                                                                                                                                                                      /organism="unknown"
1623 c 1532 g
                                                                                                                                           AR118686
Sequence 53 from patent
AR118686
AR118686.1 GI:14100596
                                                                                                                                                                                                                                                                                                                                     3.22e-67
943.00
100.008
98.358
98.958
                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                  1060 a
                                                                                Ser 182
                                                                                                   541 TCC 543
                                                                                                                                                                                              Unknown
                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                182
                                                                                                                                           LOCUS
DEFINITION
                                                                                                                                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                      AUTHORS
TITLE
JOURNAL
                                                                                                                                                                ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                             REFERENCE
                                                                                                                       RESULT 5
AR118686
                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      D,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                do
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                       ò
                                                          g
                                                                                ò
                                                                                                                                                                                                                                             /codo_start=1
/product="NS3 protease"
/product="NS3 protease"
/product="NS3 protease"
/db_xref="C1:111550.1"
/db_xref="C1:11550.1"
/translation-"APITAYAQOTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFL
ATCINGVCWTVYHGAGTRIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSS
DI;YUTWHADVIPVRRRGDSNGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCT
RGVARAVDFIPVETTMRS"
109 c 1159 g 103 t
                                Genetic diversity and response to IFN of the NS3 protease gene from clinical strains of the hepatitis C virus Arch. Virol. 147 (7), 1385-1406 (2002)
                    and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                                                       2 (bases 1 to 543)
Holland-Staley,C.A., Kovari,L.C., Golenberg,E. and Mayers,D.L.
Direct Submission
Submitted (09-ARR-2001) Infectious Disease Research, Henry Ford
Health Systems, 2799 W. Grand Blvd. Rm 7045 E & R, Detroit, MI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 (bases 1 to 543)
Holland-Staley,C.A., Kovari,L.C., Golenberg,E.M., Pobursky,K.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ThrargThrIlealaSerProLysGlyProValIleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                       543
180
0
0
                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                               1. .543
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/isolate="Pt.lV"
                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                      /db_xref="taxon:11103"
/note="type: 1A"
<1, .>543
/note="polyprotein"
                                                                                                                                                                                                                                                                                                                                                                                                                                             (1-543)
                                                                                                                                                     Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-1 (1-182) x AF369235
                                                                                                                                                                                                                                                                                                                                                                     1.89e-68
945.00
100.00%
99.45%
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity:
Query Match:
DB:
                              Mayers, D.L.
                                                                                 12111414
                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                          JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                        .. NO. .
                                                                                                             TITLE
JOURNAL
                                                                                                                                                                                                                             CDS
                                                                                                                                                      FEATURES
                                       TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
1049
                                                                                                                                                                                                                                                                                                                                                                                                                                           1230 GGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCCGGGG 1289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyAspSerArgGlySerLeuLeuSerProArgProIlcSerTyrLeuLysGlySerSer 140
                                                                                                        PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                    141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
9
                                                                                                                                                                                                                                                                                                                                                                                                              80
                                                                                                                                                                                                                                                                                                                                                                                                             41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                  5360
179
3
0
0
                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
```

us-09-965-594-1.rge

Db 1410 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGACCTAGAGACATG 1469 Qy 181 ArgSer 182	1 7	LOCUS LOCUS LOCUS LOGIS LOGIS LOGIS LOGIS ACCESSION 109328 VERSION 109328.1 G1:587963 KEYWORKS SOURCE UNKNOWN.	Unclassified. REFERENCE 1 (bases led. AUTHORS 1 (bases led.) AUTHORS Houghton, M., Choo, OK. and Kuo, C. JOURNAL Patent: WO 8904669-A 8 01-JUN-1989; PEATURES 1.5360 BASE COURT 1061 a 1623 c 1533 g 1143 t ORIGIN	3.22e-67 Length: 943.00 Matches:	Percent Similarity: 100.00% Conservative: 3	IleThralaT ATCACGGCGT	Qy 21 SerLeuthrGlyArgAspLy9AsnGlnValGluClyGluValGlnIleValSerThrAla 40	Qy 41 AlaGlnThrPheLeuAlaThrCySIleAsnGlyValCySTrpThrValTyrHiSGlyAla 60	Oy 61 GlythrargthrilealaScrProLysGlyProValileGlnMetTyrThrasnValasp 80	0y 81 LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100 :::	Oy 101 GlySerSerAspLeuTyrLeuValThrargHisAlaAspVal1leProValArgArgArg 120	Oy 121 GlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLcuLysGlySerSer 140	Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160	Qy 161 ThrArgGlyValAlaLysAlaValAspPheIJeProValGluSerLeuGluThrThrMet 180	Oy 181 ArgSer 182
ThrargGl 	Oy 181 ArgSer 182 1 1 1 1470 AGGTCC 1475	RESULT 6 106434 106434 5360 bp DNA linear PAT 02-DEC-1994 DEFINITION Sequence 48 from Patent EP 0318216. ACCESSION 106434 TENSION 106434.1 GI:590311	KEYMORDS SOURCE ORCANISM Unknown. ORCANISM Unknown. Unclassified. REFERENCE 1 (bases 1 to 5360) AUTHORS Houghton, Choo, 0L. and Kuo, G. TITLE Nanby diagnostics and vaccines JOURNAL Patent: EP 0318216-A1 48 31-MAY-1989; FEATURES. Location/Qualifiers	ט	Alignment Scores: 3.22e-67 Length: 5360 Pred. No.: 943.00 Matches: 179 Percent Similarity: 100.00% Conservative: 3 Best Local Similarity: 98.35% Mismatches: 0	98.95% Indels: 6 Gaps: 2) x IO6434 (1-5360)	Oy 1 MetalaprollethralatyralaglnglnthrargglyLeuLeuglyCysIleIlethr 20 :::1))	Qy 21 SerLeuthrGlyArgAspLysasnGlnValGluGlyGluValGlnIleValSerThrAla 40	Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60 	Qy 61 GlyThrArgThrileAlaSerProLysGlyProValileGinMetTyrThrAsnValAsp 80 [11] [1] [1] [1] [1] [1] [1] [1] [1] [1	Oy 81 LysAspLeuValGlyfTpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100 :::	Oy 101 GlySerSerAspLeuTyrLeuValThrargHisAlaAspVallleProValArgArgArg 120 11 11 11 11 11 11 11	Oy 121 GlyaspSerArgGlySerLeubeuSerProArgProlleSerTyrLeuLySGlySerSer 140	Oy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyllePhcArgAlaAiaValCys 160 (1111111111111111111111111111111111	Qy 161 ThrangGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180

us-09-965-594-1.rge

```
US-09-965-594-1 (1-182) x I06440 (1-6785)
                                                                                                  Location/Qualifiers
1. .6785
                                                                                                                /organism-"unknown"
1 2050 c 1914 g
      Sequence 54 from Patent 106440
                                                                                                                                                        4.15e-67
943.00
100.00%
98.35%
98.95%
                          GI:590312
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1743 AGGTCC 1748
                                                                                                                                                                 Score:
Percent Similarity:
1
Rest Local Similarity:
9
Ouery Match:
6
                                                         Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 ArgSer 182
                                                                                                                        1392 a
                         106440.1
                                        Unknown
                                                  Unknown
 106440
                                                                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                61
                                                                                                           source
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                        BASE COUNT
ORIGIN
                                                                         AUTHORS
TITLE
JOURNAL
                                                ORGANISM
                                                                  REFERENCE
                                                                                                  FEATURES
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                              \dot{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S
D
                                                                                                                                                                                                                                                       1562
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1623 GGGGGTCCGCTGTTGTGCCCCGCGGGCACGCCGTGGGCATATTTAGGGCCGGCGGTGTGC 1682
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     141 GlyGlyProLeuLeuCySProAlaGlyHisAlaValGLyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                      101 GlyserSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
                                                                                                                                                                                                                                                                                                                       41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIrpThrValTyrHisGlyAla 60
                                                                                                                                                                                                                                                                                                                                                                                                                               ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet.
                                                                                                                                                                                                 m000
                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                    Unknown.
Unclassified.
Unclassified.
I (bases 1 to 6785)
Chien, D.Y.
NANBV diagnostics and vaccines
Location/qualifiers
Location/qualifiers
1.6785
Location/qualifiers
1.6785
                          DN'
                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                        US-09-965-594-1 (1-182) x AR118692 (1-6785)
                                 patent
                        AR118692
Sequence 65 from patent
AR118692
AR118692.1 GI:14100562
                                                                                                                                                                               4.15e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AGGTCC 1748
                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                 Unknown.
                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                        1323 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                        LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                  source
                                                                 SOURCE
ORGANISM
                                                                                                                                                 BASE COUNT
                                                                                                                                                                                   ..
Q
                                                                                                  AUTHORS
TITLE
                                                                                         REFERENCE
                                                                                                                 JOURNAL
        RESULT 8
AR118692
                                                                                                                          FEATURES
                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     요
```

```
1382
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1562
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAT 02-DEC-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100
 02-DEC-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                      80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1323 GCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGGGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                               1383 GGAACGAGGACCATGGCGTCACCGAAGGTCCTGTCATCCAGATGTATACCAATGTAGAC
                                                                                                                                                                                                                                                                                                                                                                                           41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                     GlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
 linear
                                                                                                                                                                                                         6785
179
3
0
0
                                                                                                                                                                                                                   Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                    1 (bases 1 to 6785)
Houghton, M. Choo, Q.-L. and Kuo, G.
Nanby diagnostics and vaccines
Patent: EP 0318216-A1 54 31-MAY-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS 109329 6785 bp
DEFINITION Sequence 10 from Patent WO 8904669.
6785 bp
EP 0318216.
```

Page

```
Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown
                                                                              Alignment Scores:
                                                                                                                                                                                                                                           1908
                                                                                                                                                                                                                                                                                                2028
                                                                                                                                                                                                                                                                                                                                                    2148
                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                             121
                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
I09331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                     Pred. No.:
ORGANISM
            REFERENCE
AUTHORS
TITLE
JOURNAL
                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                    1,000
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                ò
                                                                                                                                                           qq
                                                                                                                                                                         ò
                                                                                                                                                                                    q
                                                                                                                                                                                                 õ
                                                                                                                                                                                                               g
                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                            οχ
                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                     PAT 16-MAY-2001
                                                                                                                                                                   81 LysaspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100
                                                                                                                                                                                                                                                                                                                                                           GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                  80
                                                                                                                                                                                                                                                  GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMeLTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                        6785
179
3
0
0
                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         DNA
                                Unclassified.
1 (bases I to 6785)
Houghton,M., Choo,Q.-K. and Kuo,G.
Patent: WO 8904669-A 10 01-JUN-1989;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                     7310 bp
Sequence 74 from patent US 6150087.
AR118696.
AR118696.1 GI:14100606
                                                                              1429
                                                                                                                                         Gaps:
                                                                                                                                                      US-09-965-594-1 (1-182) x I09329 (1-6785)
                                                                      /organism="unknown"
2050 c 1914 g
                                                                                                        4.15c-67
943.00
100.00%
98.35%
98.95%
      GI:587964
                                                                                                                                                                                                                                                                                                                                                                                                                        111111
AGGTCC 1748
                                                                                                                                                                                                                                                                                                                                                                                                                 181 ArgSer 182
                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                               ۵
     109329.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown
                  Unknown
                          Unknown
                                                                                                   Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                             1743
                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 11
AR118696
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                    Query Match:
DB:
                                                                  source
                           ORGANISM
                                             AUTHORS
JOURNAL
FEATURES
                                                                              BASE COUNT
                                                                                                   Alignment (
Pred. No.:
ACCESSION
                                        REFERENCE
      VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SOURCE
                    SOURCE
                                                                                     DRIGIN
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                       a
                                                                                                                                                                                                                                              6 G 6
                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                qq
                                                                                                                                                                                            ò
                                                                                                                                                                                                                        8 6
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                    ò
```

```
2147
                                                                                                                                                                                                                                               1847
                                                                                                                                                                                                                                                                                                                          1961
                                                                                                                                                                                                                                                                                                                                                               2027
                                                                                                                                                                                                                                                                                                                                                                                                   2087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAT 02-DEC-1994
                                                                                                                                                                                                                                                                                      1907
                                                                                                                                                                                                                                                                                                                                                                                                                                                            141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                                                                                                                                                                       140
                                                                                                                                                                                                                                                                                                                                              100
                                                                                                                                                                                                                                                                    09
                                                                                                                                                                                                                                                                                                         80
                                                                                                                                                                                                                              40
                                                                                                                                                                                                                                                                                                                                                                                 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                  61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                       Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7310 bp
Sequence 15 from Patent WO 8904669.
109331.1 GI:587966
                                                                                                                                           Indels:
Gaps:
                                                                                                                                                                        US-09-965-594-1 (1-182) x AR118696 (1-7310)
1 (bases 1 to 7310)
Chien.D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 74 21-NOV-
Location/Qualifiers
                                                       /organism="unknown"
1495 a 2220 c 2056 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unclassified.
1 (bases 1 to 7310)
                                                                                                      4.5e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2268 AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 ArgSer 182
                                                                                                                        Percent Similarity:
Best Local Similarity:
```

```
TYMOYLPOESTTLEALISTATION WOONTHEAGENCULEDINGKEELED TOTOWOYLPOEST WOONTHEAGENCLIED TOTOWOYLPOEST WAS CONTRICTED WITH FULL TOTOWAY TO THE ACT OF 
                                                                                                                                                                                                                                                                                                                                                         /translation="GCPERLASCRPLTDFDQQWGPISYANGSGPDQRPYCWHYPPRPC
GIVPAKSVCGPVYCFTPSPVVVGTTDRSGAPTYSWGENDTDVFVLNNTRPPLGNWFGC
TWMNSTGFTKVCGAPPCVIGGAGNNTLHCPTDCFRKHPDATYSRCGSGPWITPRCLVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QEMGGNITRVESENKVVILDSPDPLVAEEDEREISVPAEILERGRRFAQALPVWARPD
YNPPLVETWKKPDYEPPVVHGCPLPPPKSPPVPPRKKRTVVLTESTLSTALAELATR
SFGSSSTSGITGDNTTISSEPAPSGCPPDSDAESYSSMPPLEGEPGDPDLSDGSWSTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDLGVRVCEKMALYDVVTKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKKTPMGFSYD
TRCFDSTVTESDIRTEEAIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENGGYRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RAFTEAMTRYSAPPGDPPOPEYDLELLTSCSSNVSVAHDGAGKRVYYLTRDPTTPLAR
AAWETARHTPVNSWLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALDCEIYGACY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       QRQKKVTFDRLQVLDSHYQDVLKEVXAAASKVKANLLSVEBACSLTPPHSAKSKFGYG
AKDVRCHARKAVTHINSVWKDLLEDNVTPIDTTIMAKNEVFCVQPEKGGRKPARLIVF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SSEANAEDVVCCSMSYSWTGALVTPCAAEEQKLPINALSNSLLRHHNLVYSTTSRSAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CRASGVLTTSCGNTLTCYIKARAACRAAGLQDCTMLVCGDDLVVICESAGVQEDAASL
                                                                                                                                                                                                                                                                                                                                                                                                                           YPYRLWHYPCTINYTIFKIRMYVGGVEHRLEAACNWTRGERCDLEDRDRSELSPLLLT
      Original source text: Hepatitis C virus, cDNA to viral RNA, clones K9-1 through 15e, isolated from chimpanzee (individual 910) blood plasma.
Draft entry and printed sequence for [1] kindly submitted by M.Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street, Emeryville CA 94608.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7310
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                               /protein_id="AAA45677.1"
/db_xref="G1:329876"
                                                                                                                                                              1. .7310
/organism="Hepatitis C ./mol_type="genomic RNA"
/db_xref="taxon:11103"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-1 (1-182) x HPCPOLYP (1-7310)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2058 9
                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                      <1. .>7310
/note="polyprotein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIEPLDLPPIIQRL
                                                                                                                                                                                                                                                                                                 /codon_start=3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.5e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2218 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                   CDS
                                                                                                                                          FEATURES
         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           02-AJG-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallieProValArgArgArg 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        141 GlyGlyProLeuCeuCysProAlaGlyHisAlaVaiGlyflcPheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2148 GGGGGTCCGCTGTTGTGTGTCCCCGCGGGGCACGCCGTGGGCATATTTAGGCCGCGCGGGGTGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 (bases 1 to 7310)
Choo.O.-L., Richman,K. and Han,J.
The nucleotide sequence of the Hepatitis C viral genome
Unpublished (1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus polyprotein gene, partial cds. M32084.1 G1:329875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         no DNA
                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses,
Hepacivirus.
Houghton,M., Choo,Q.-K. and Kuo,G.
Patent: WO 8904669-A 15 01-JUN-1989;
Location/Qualifiers
                                                                                                                   1539
                                                                                                                                                                                                                                                                                                               Caps:
                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-1 (1-182) x 109331 (1-7310)
                                                                                            /organism="unknown"
2218 c 2058 g
                                                                                                                                                                                                     4.5c-67
943.0c
100.00%
98.35%
98.95%
                                                                         .7310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          111111
2268 AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 Argser 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    polyprotein
                                                                                                                                                                                                                      Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                     ø
                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HPCPOLYP
LOCUS
DEFINITION
ACCESSION
VERSION
                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KEYWORDS
SOURCE
ORGANISM
                                                                                                                   BASE COUNT
                                                                                                                                                                                                     No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
JOURNAL
                         JOURNAL
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
                                                  FEATURES
                                                                                                                                       ORIGIN
                                                                                                                                                                                                     Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                · &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ç
```

MetalaProlleThralaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysllelleThr 20 ò

2973

80

100

2913

09

3093

140

3153

160

3213

180

3273

qq

ò q οy

ò

đ S a В ò Op οχ qq ò 윤 δ g

ογ

```
PAT 16-MAY-2001
                                                                                                                            161 ThrargGlyValalaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet
                                                                                                                                                                                                                                    LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                              GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8987
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                           DNA
                                                                                                                                                                                                                                                                                                                                                                         1 (bases 1 to 8987)
Chien.D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 137 21-NOV-2000;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                           7 bp DN.
6150087.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-1 (1-182) x AR118728 (1-8987)
                                                                                                                                                                                                                                                                                                                                                                                                                         /organism="unknown"
2735 c 2547 g
                                                                                                                                                                                                                                                                                                          AR118728 8987
Sequence 137 from patent US
AR118728
AR118728.1 GI:14100638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5.62e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                   8987
                                                                                                                                                                                                                                                                     3274 AGGTCC 3279
                                                                                                                                                                                                                                                                                                                                                                   Unclassified.
                                                                                                                                                                                                                                                           181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                    Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                   1807
                                                                                                                                                                                                                                                                                                                                                   Unknown
                                                                                                                                                                                                                                                                                                                                                            Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
Pred. No.:
                                                                                             81
                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                  DEFINITION
ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                   JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                          RESULT 15
AR118728
LOCUS
                                                                                                                                                                                                                                                                                                                                          KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                          d
            q
                           õ
                                           d
                                                             ò
                                                                           qq
                                                                                             ò
                                                                                                            qq
                                                                                                                           õ
                                                                                                                                           g
                                                                                                                                                             ŝ
                                                                                                                                                                          g
                                                                                                                                                                                           ^{\circ}
                                                                                                                                                                                                          Db
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                          Ω̈́Ω
                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2027
                                                                                                                                                                     2087
                                                                                                                                                                                                    2:47
                                                                                                                                                                                                                                    2207
                                                                                                                                                                                                                                                   1788 AGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT 1847
                                                                                                                                                                                                                                                                                                                                    PAT 16-MAY-2001
                                                                                                                                                                                     140
                                                                                                                                                                                                                    160
       CIGGCGCCCAICACGCCGIACGCCCAGCACAAGGGGCCTCCIAGGGIGCATAAICACC 1787
                                                                                                                      100
                                                                                                                                                      GlySerScrAspLcuTyrLcuValThrArgHisAlaAspValIleProValArgArgArg 120
                       40
                                                                                                                      SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8316
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                     DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ĻJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                   1 (bases 1 to 8316)
Chieb,D.Y.
MANDY diagnostics and vaccines
Patent: US 6150087-A 88 21-NOV-2000;
                                                                                                                                                                                                                                                                                                                               Sequence 88 from patent US 6150087.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JS-09-965-594-1 (1-182) x AR118703 (1-8316)
                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                   /organism="unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                            2345 g
                                                                                                                                                                                                                                                                                                                                                            GI:14100613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5.17e-67
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                           2529 c
                                                                                                                                                                                                                                                                                                                                                                                                                                          8316
                                                                                                                                                                                                                                                                                              Unclassified.
                                                                                                                                                                                                                                                                                    181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                           AR118703.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                            1671
                                                                                                                                                                                                                                                                                                                                                                            Unknown.
                                                                                                                                                                                                                                                                                                                                                                                     Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                    2268
                         21
                                                                                                                                                      101
                                                                                                                                                                                                                                                                                                                  RESULT 14
AR118703
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                           SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                            AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score:
```

ò

```
3375
                   141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                Argser 182
                                    3496
                                              181
      S G
$ Q
```

Search completed: August 31, 2003, 00:45:16 Job time : 2382.6 secs

```
WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAA70344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
HCV-1 NS3/4a confo
HCV genomic amino
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence encoded i
Protein sequence o
Peptide encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV-1 NS3/4a mutan
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                     August 30, 2003, 17:42:58 ; Search time 41.2251 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A_Geneseq_19Jun03:*

| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:*
| SIDSI/gcgdata/geneseqy-embl/AA1999.DAT:*
| SIDSI
                                                                                                                                                                                                                                                                                                                           1 MAPITAYAQQTRGLLGCIIT......GVAKAVDFIPVESLETIMRS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first, 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAE18689
AAU76377
ABG72261
AAR40120
AAR51170
AAP92041
AAP90158
                                                                                                             OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15211
                                                                                                                                                                                                                                                                                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB
                                                                                                                                                                                                                                                                        US-09-965-594-1
953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               %
Ouery
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3011
609
1766
1786
1786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0
99.3
99.3
99.3
99.0
99.0
99.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sed
                                                                                                                                                                                                                                                                           Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
DB
                                                                                                                                                                                                                                                                                                                              Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum Maximum
                                                                                                                                                                Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Š
```

				,		,			
	4 4 4 4 4 4 4 4	υ 0 υ 0		2301	2 5	AAP92047			encoded
٦ ,		n c		, ,	9 6	AAF 92030		, '	
v (n 0	> <	ית קיי	2 6	AAP90288			
.	•	a	٥.	7.7	21	AAB18540		_	encoded by
4	•	S)	0.	8	14	AAR34009		HCV-1 poly	protein
2	₹.	σ.	0.	8	16	AAR70230		Composite	ite hepatiti
9	₹	On.	0.	95	50	AAY14975		Amino acid	sedneuc
7	₹	a)	0.	95	21	AAB18541		Polyprotein	encode
80	₹.	ð	0.	2	13	AAR21519		Compiled HCV	
σ,	₩.	on i	0.	5	4	AAR31621		Hepatitis (C virus
0	4	on .	0.	5	17	AAR90931		Hepatitis (virus
- 4:	ᢦ ⋅	on a	0.	5	18	AAW34480		HCV polyprotein	tein.
~1	σ.	n e	0.	5	13	AAW4 0038		н	tein.
	•	on a	٥.	51	53	AAE22049		Hepatitis (C virus
.	•	n o	٥,	55	53	AAU84597		HCV polypro	tein 1
n 4	• •	x 0 0	20 , c	632	7 6	AAE21847		Hepatitis	virus
o r	•	o o	.	200	7 7	AAE19905		Hepatitis	virus
~ 0	•	xo c	20 C	989	5 6	AAE21837		Hepatitis	virus
o d	* 4	o a	0.0	900	7 0	AAE21838		Hepatitis	virus
r C	* 4	oα	oα	9 00 00	ה ה	AAF21840		Honatitie	SULTA Y
· ·	. 4	0		989	23	AAE21841		Hebatitis C viru	Virus
1 (2)	4	œ	ω.	686	23	AAE21842		Hepatitis	virus
3	**	œ	8.	989	23	AAE21843		Hepatitis	virus
4	4	æ	æ.	989	23	AAE21844		Hepatitis	virus
Ŋ	4	æ	ω.	989	23	AAE21845		Hepatitis	virus
و	₹	æ	∞.	989	23	AAE21846		Hepatitis	virus
7	942	8	ω.	989	23	990		Hepatitis (virus
œ ·	4	∞ .	œ.	989	23	960		Hepatitis (virus
on (ჟ∙	8	۰.	989	23	990		Hepatitis	virus
۰,	4.	o o	20.0	686	23	166		Hepatitis	virus
_	4	σ.	ω.	989	53	992		Hepatitis	virus
(7)	ͺ	86	۰.	989	23	992		Hepatitis	virus
-n -	4.	20 0	æ. (989	23	992		Hepatitis	virus
4	745	χ. Σ	æ.	989	53	266		Hepatitis	virus
	₹	8	∞.	989	23	AAE19924		Hepatitis	virus
						ALIGNMENTS			
RESULT 1									
<u> </u>	П	standard	ard;	protein	In;	182 AA.			
AX AC AAB1521	11;								
19-DEC	-2000		(first	t entry	γ,				
DE Hepatiti	tis C		virus	NS3 pr	protease	se.			
Hepati	:, 1	; NS3		protease; viral liver cancer.	vir cer.	replication;	chronic 1	liver disease	••
OS Hepatitis	tis C		virus.						

```
XX AAB15211;
XX

AAB15211;
XX

19-DEC-2000 (first entry)

E Hepatitis C virus NS3 protease.

XX

WA Hepatitis; NS3 protease; viral replication; chronic liver disease;

XX

WA WO200040707-A1.

XX

WA WO200040707-A1.

XX

WA WA 13-JUL-2000.

XX

WA WA 13-JUL-2000.

XX

WA WA 13-JUL-2000.

XX

WA WA 1999; 99US-0115271.

XX

CBRIM ) BRISTOL-MYERS SQUIBB CO.

XX

WA WPI; 2000-465976/40.

XX

WPI; 2000-465976/40.

XX

WPI; 2000-465976/40.

XX

WA WO 1500-465976/40.

XX

WA WA 12000-465976/40.

XX

WA WA 12000-465976/40.

XX

WO 15000-465976/40.

XX

WO 15000-465976/40.

XX

WO 15000-465976/40.

XX

WO 15000-465976/40.

YA WO 150
```

a

```
N-PSDB; AAD29795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 RS 182
  Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200196870-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  - S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
  TXX BD X X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B A X B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ó.
                                                                                                                 The present sequence is the Hepatitis C virus (HCV) NS3 protease enzyme. This protein is essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. NS4A is also needed for this process and inhibitors of the two proteins should act as antivitral treatments of ACV infection. This is useful as MCV can lead to chronic liver disease such as cirrhosis. Invertail as MCV can lead to chronic liver disease such as cirrhosis. Invertail and NS3 MS4A fission proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease.inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSIJTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MAPITAYAQQTRGLLGCI.TSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MAPITAYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus: NS3/4a antigen; HCV infection; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Coit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0
                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 182;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tandeske L, George-Nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note= "Wild type Thr substituted with Pro"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Wild type Ser substituted with Ile"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 953; DB 21;
Pred. No. 3.9e-91;
: Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV-1 NS3/4a mutant conformational antigen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAE18689 standard; Protein; 586 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0
                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0%;
                                                                        Claim 3; Fig 9; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14-JUN-2001; 2001WO-US19369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 100.
Best Local Similarity 100.
Matches 182; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus type 1. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                               182 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200196875-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Chien DY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAE18689;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
61 GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present invention relates to hepatitis C virus (HCV) core antigen and NS (nonstructural) 3.4a antibody combination assay that can detect both HCV antigens and antibodies present in a sample using a single solid matrix as well as immunoassay solid supports for use in the assay. The solid support is useful for detecting HCV infection in a biological sample. The present sequence is HCV-1 NS3/4a mutant conformational antigen. This sequence is used in the exemplification of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion; immunoassay solid support; multiple epitope fusion antigen; MEFA; non-structural protein; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus NS3/4a conformational epitope protein sequence
                                                                                                                                 Immunoassay solid support useful for detecting hepatitis C virus infection in a biological sample, comprises at least one of HCV anti-core antibody and HCV NS3/4a epitope, bound to the support
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Wild-type Thr substituted by Pro"
Misc-difference 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 946; DB 23; Length 6 Pred. No. 1.2e-89; 2; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              by Ile"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              substituted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         *Wild-type Ser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAU76377 standard; Protein; 686 AA.
                                                                                                                                                                                                                                                                                                                       Example 2; Fig 4; 87pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 98.9%;
Matches 180; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14-JUN-2001; 2001WO-US19156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-JUN-2000; 2000US-212082P-02-APR-2001; 2001US-280811P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note-
2002-179522/23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Misc-difference 403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 686 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAR X PAR X
```

ö

```
MEFA; anti-HCV antibody; NS3/4a conformational antigen;
HCV infection; mutant; mutein.
                                                                                                                                                                                                                                                                            15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                        GEORGE-NASCIMENTO C.
                                                                                                                                                                                                                                                   14-JUN-2001; 2001US-0881654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 98.99
Matches 180; Conservative
                                    Hepatitis C virus type 1. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                  MEDI/) MEDINA-SELBY A.
                                                                                                                                                                                                                                                                                                                                                                                                                            Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2003-147573/14.
N-PSDB; ABX14410.
                                                                                                                                                                                                                                                                                                                                             ARCANGEL P.
TANDESKE L.
                                                                                                                                 Misc-difference 403
                                                                                                                                                         Misc-difference 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     686 AA;
                                                                                                                                                                                                                                                                                                                                  CHIEN D Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                        Medina-Selby A;
                                                                                                                                                                                                US2002146685-A1
                                                                                                                                                                                                                                                                                                                                                                                    Ö.
                                                                                                                                                                                                                           10-0CT-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                             Chien DY,
                                                                                                                                                                                                                                                                                                                                                                        (CEORY)
                                                                                                                                                                                                                                                                                                                                  CHIE/
                                                                                                                                                                                                                                                                                                                                                ARCA/)
                                                                              Key
Reqion
                                                                                                                                                                                                                                                                                                                                                           TAND/)
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention relates to a new immunoassay solid support consisting essentially of at least one hepatitis C virus (HCV) NS3/4a conformational epitope and a multiple epitope fusion antigen (MEFA), bound to the support. The NS3/4a conformational epitope and/or MEFA reacts specifically with anti-HCV antibodies prosent in a biological sample from an HCV-infected individual. The immunoassay of the invention is useful for detecting hepatitis C virus infection in a biological sample. The method of the invention provides a sensitive, accurate diagnostic and prognostic tool to provide adequate patient care and to prevent transmission of MCV by blood and by blood products, or by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                    prevent transmission of may by plood and by blood products, or by personal contact. Use of NS3/4a conformational epitope in combination with MEFA, provides a sensitive and reliable method for detecting early HCV seroconversion. Use of MEFA has the added advantages of decreasing masking problems, improving sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit surface area of substrate, and improving substrate. Detection accuracy is increased and the incidence of false results is reduced because of the identification and the use of highly immunogenic HCV antigons which are present during the early stages of HCV seroconversion. The present amino acid sequence represents the non-structural protein NS3/4a conformational epitope of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MAPITAYAQQTRGI.LGC11TSLTGRDKNQVFGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Gaps
                                                                                                                                Immunoassay solid support, useful for detecting hepatitis C virus infection in biological sample, comprises HCV NS3/4a conformational epitope and multiple epitope fusion antigen bound to the support \cdot
                                                      Coit D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Immunoassay solid support; Hepatitis C Virus type-1; HCV-1;
NS3/4a conformational epitope; multiple epitope fusion antigen;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 946; DB 23; Length 686;
Pred. No. 1.2e-89;
2; Mismatches 0; Indels
                                                      Tandeske L, George-nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV-1 NS3/4a conformational antigen.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABG72261 standard; Protein; 686 AA.
                                                                                                                                                                                     Claim 5; Fig 3; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99.3%;
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-MAR-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              180; Conservative
                                                     Chien DY, Arcangel P, Medina-selby A:
                                                                                            WPI: 2002-090228/12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 180; Conserv
                            (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          586 AA;
                                                                                                         N-PSDB; ABK15344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABG72261;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ωp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
```

```
The present invention relates to immunoassays comprising Hepatitis C Virus (HCV) NS3/da conformational epitope and multiple epitope fusion antigen (MEFA), bound to a solid support. The NS3/da epitope and/or the multiple epitope fusion antigen react with anti-HCV antibodies present in a biological sample from an HCV-infected Individual. The immunoassays and methods of the invention are useful for detecting they infected in a biological sample. The Inventive immunoassay solid support provides a sensitive and reliable method for detecting early HCV seroconversion. The assays can detect HCV infection acused by any C six known genotypes of HCV as says can detect the multiple epitope fusion proteins decreases masking problems, improves sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit area of substrate, and improves selectivity. The present sequence (C represents HCV type 1 (HCV-1) NS3/da conformational antigen, a mutant of the HCV-1 NS3/da polypeptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 HAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Immunoassay solid support for detecting Hepatitis C Virus infection in biological samples, comprises Hepatitis C Virus conformational epitope and multiple epitope fusion antigen -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MAPITAYAQQIRGLIGCIITSLIGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                         /note "Corresponds to amino acid residues 1027-1711 of HCV-1 NS3/4m polypeptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Coit D;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                /note- "Substitution of wild-type Ser to Ile"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 686;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             George-Nascimento C,
                                                                                                                                                            /note- "Substitution of wild-type Thr to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 946; DB 24;
Pred. No. 1.2e-89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99.3%; Sco. 98.3%; Pred. No. 1.-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Tandeske L,
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 2; Fig 3A-3D; 45pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 op.
```

g

g ò

```
GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide fragments of the non-structural protein (NS3) are reactive with and can detect antibodies against the NS3 domain of HCV. The peptides can be used for diagnosis of HCV infection. Nonspecific reaction can be inhibited and misdiagnosis of HCV infection can be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New peptide(s) reactive with anti-hepatitis C virus antibody for specific, early diagnosis of HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 609;
                                                                                                                                                                                                                                                                                                                                Peptide; antibody; hepatitis C virus; HCV; identification; diagnosis; non-A non-B hepatitis; NANB; detection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99.0%; Score 943; DB 15; 198.4%; Pred. No. 2.1e-89; iive 3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3; Mismatches
                                                                                                                                                                                                                                                                                                Hepatitis C virus non-structural protein 3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Page 9-10; 15pp; Japanese.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAP92041
ID AAP92041 standard; protein; 1766 AA.
                                                                                                                                                                                         AAR51170 standard; peptide; 609 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (OLYU ) OLYMPUS OPTICAL CO LTD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           92JP-0209201.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92JP-0209201.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      decreased. See AAR51162-70
                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 98.4*
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1994-106803/13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       609 AA;
                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
                                                                                          ||
| RS 1207
                                                                    RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          05-AUG-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           05-AUG-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAR-1994.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RS –
                                                                                                                                                                                                                                                             20-0CT-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                          AAR51170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence
                                                                    181
                                                                                                                                                          RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 7
                                                                                                                                                                         AAR51170
                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                               a
                                                                    \tilde{c}
                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTRT1ASPKGPV1QMYTNVIJKDLVGWPAPQGSRSI.TPCTCGSSDLYI.VTRHADVIPVRRR 120
                               GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RNA was isolated from the plasma of a HCV seropositive human (designated "LG") and cDNA was prepared from it. The cDNA was PCR amplified using specific primers with sequences based on the prototype HCV-1 DNA sequence (GENBANK M62321). Further amplification using nested primers resulted in 7 adjacent HCV DNA fragments which could be assembled into a full-length sequence. The DNA sequence. Comparison of the LG genomic amino acid sequence. Comparison of the LG genomic amino acid sequence with that from HCV-1 showed 134 amino acid differences.
                                                                                                                                                                                                                                                                                                                                HCV genomic amino acid sequence isolated from infected human LG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New plasmid pHCV-162 is a mammalian expression systems for HCVI proteins - useful for diagnosing HCV infection and as vaccines for preventing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                      Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV; NANBHV; human growth hormone; HGH; secretion signal; fusion protein; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Frail
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99.1%; Score 944; DB 14; ilarity 98.4%; Pred. No. 1.4c-88; Conservative 3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Devare SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 1; Page 39-49; 100pp; English.
                                                                                                                                                                                                            AAR40120 standard; Protein; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Desai SM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93WO-US00907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92US-0830024
                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Zeck BJ;
                                                                                                                                                                                                                                                                                (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1993-258673/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Casey JM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (ABBO ) ABBOTT LAB
                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C Virus.
                                                                                                                       182
                                                                                      RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Yamaguchi J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MO9315193-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29-JAN-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-JAN-1992;
                                                                                                                                                                                                                                                                              25-MAR-2003
                                                                                                                                                                                                                                                                                                27-JAN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           05-AUG-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                    181
                                                                                                                                                                                                                                           AAR40120;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bode SL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                         RESULT 5
```

qq à 윱

8

ô

us-09-965-594-1.rag

```
AAP90158 standard; protein; 1786 AA
                                    (updated)
(first entry)
                                                                                                                                                                                                                                   Houghton M, Choo QL,
                                                                                                                                                                                                                                                    WPI; 1989-215054/30
                                                                                                                                                                                                                                                                                                                                                                                       1786 AA;
                                                                                                                                                                                                                                                              N-PSDB; AAN90327
                                                                                                 Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RS 491
                                                                                                                                                     18-NOV-1988;
                                                                                                                                                                      18-NOV-1987;
30-DEC-1987;
                                                                                                                                                                                       26-FEB-1988;
26-OCT-1988;
                                   25-MAR-2003
10-NOV-1989
                                                                                                                  GB2212511-A.
                                                                                                                                   26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      490
                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                  AAP90158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
AAP90164
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           It is the sequence encoded in the open reading frame of hepatitis C virus cDNA inserts in clones 14i,m 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, a33f, 33g and 39c. It is antipenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
                                                Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 141, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 8i, 32, 33b, 25c, 14c, 8f. 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                   Claim 13; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 943; DB 10; Length 1766;
Pred. No. 8.8c-89;
3; Mismatches 0; Indels 0
                                                                                       Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                 and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                       (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                             870S-0132714.
870S-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                             Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   99.08;
                                                                                                                                                              88EP-0310922
                           (updated)
(first entry)
                                                                                                                                                                                                                                                                                                       Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 98.4
Matches 179; Conservative
                                                                                                                                                                                                                                                           Houghton M, Choo QL,
                                                                                                                                                                                                                                                                             WPI; 1989-159274/22
                                                                                                                                                                                                                                           (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                 1766 AA;
                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                        N-PSDB: AAN92097
                                                                                                                                                                                                                                                                                                                                                                                               non-B hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |-|
RS 491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS 182
                                                                                                                                                                                                26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
                                                                                                                                                              18-NOV-1988;
                                                                                                                                                                               18-NOV-1987;
30-DEC-1987;
                          25-MAR-2003
02-MAR-1990
                                                                                                                                            31-MAY-1989
                                                                                                                                                                                                                         14-NOV-1988
                                                                                                                          EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                   Sedneuce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              430
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 8
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The sequence is encoded by the composite cDNA of AAN90327. These antigens react with antibodies in patients with non-A non-B hepatitis (NANBH). They can be used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct PR [ield.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    prodn. of polynucleotide probes, dlagnosis, prevention and treatment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 943; DB 10; Length 1786;
Pred. No. 8.9e-89;
3; Mismatches 0; Indels 0
Protein sequence of hepatitis c virus composite cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP90164 standard; protein; 2261 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus gene - used for polypeptide(s) and antibodies for of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; fig 26; 30pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                          Kuo G;
                                                                                                                                                                                                                                                               87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 99.0%;
Best Local Similarity 98.4%;
Matches 179; Conservative
                                                                                                                                                                                                                       88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                           (CHIR ) CHIRON CORPORATION
                                          Hepatitis C virus; vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-MAR-2003 (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DXYX5
```

ö

9

180

15:58:50

31

Sun Aug

```
Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAP92050 standard; protein; 2436 AA.
                                                                                                                                                                                 ΰ
                                                                                          87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                           88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (updated)
                                                                                                                                                                                 Choo OL,
                                                                                                                                                                                                   WPI; 1989-159274/22
                                                                                                                                                               (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                        2301 AA;
                    Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           K9-1 through 15e
                                                                                                                                                                                                            N-PSDB; AAN92103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RS 182
                                                                                                     30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003
02-MAR-1990
                                                                                                                                  25-OCT-1988;
14-NCV-1988;
                                                                          18-NOV-1988;
                                                                                                                                                                                 Houghton M,
                                                        31-MAY-1989
                                     EP318216-A
                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             581
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAP92050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAP92050
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                  GTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 12f through 15c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                       The sequence is the peptide encoded by the composite hepatitis C virus (HCV) CDNA of AAN90331. The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, our to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                 99.0%; Score 943; DB 10; Length 2261; llarity 98.4%; Pred. No. 1.2e-88; Conservative 3; Mismatches 0; Indels 0
                                     Hepatitis C virus; clone 12f; clone 15e; probe; vaccine
                    Peptide encoded by composite hepatitis C virus cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ź
                                                                                                                                                                                                                                                                                       Disclosure; fig 32; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP92047 standard; protein; 2301
                                                                                                                                 87US-0132714.
87US-0139886.
88US-0161072.
                                                                                                                88GB-0C27024
                                                                                                                                                              880S-0263584
                                                                                                                                                                                 (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
 (first entry)
                                                                                                                                                                                                  Choo QL,
                                                                                                                                                                                                                     WPI; 1989-215054/30.
                                                                                                                                                                                                                                                                                                                                                                 2261 AA;
                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity
les 179; Conserv
                                                                                                                                                                                                                                N-PSDB; AAN90331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RS 182
                                                         Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RS-
                                                                                                                18-NOV-1988;
                                                                                                                                   18-NOV-1987;
                                                                                                                                            30-DEC-1987;
26-FEB-1988;
                                                                                                                                                              26-OCT-1988;
                                                                                                                                                                                                   Houghton M,
01-NOV-1989
                                                                          GB2212511-A
                                                                                            26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP92047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                              Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP92047

ID AAP8

XX

XX

AX

DT 25-P

DT 02-P

XX

DE SEGU
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
It is the sequence encoded in the open reading frame of hepatitis C virus (HCV) cDNA inserts in clones 12f through 15e. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.

(Updated on 25-WAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 2301;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
Purified hepatitis C virus
- and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 99.0%; Score 943; DB 10;
Best Local Similarity 98.4%; Pred. No. 1.2e-88;
Matches 179; Conservative 3; Mismatches 0;
                                                                                  Claim 13; Figure 32-1 - 32-7; 139 pp; English.
```

ö

```
GIRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The sequence is the peptide encoded by the composite hepatitis C virus (HCV) cDNA of AAN90336. The polypeptides are used to diagnose HCV-induced NARBH, to raise antibodies for immunoassay or treatment, or to produce vaccines.

(N.B. This record was resubmitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protein encoded by a cDNA compiled Hepatitis C virus cDNA clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus; HCV; antisense polynucleotide; polyprotein; viral infectivity; viral replication.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match
Local Similarity 98.4%; Pred. No. 1.3e-88;
les 179; Conservative 3; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; fig 47-1 to 47-8; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB18540 standard; Protein; 2772 AA
                                                                                                                                                                                          87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                          Kuo G;
                                                                                                                                                                                                                                                                                                                      CHIR ) CHIRON CORPORATION.
                                                                                                                                                88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     15-JAN-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                     Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1989-215054/30.
N-PSDB; AAN90336.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2436 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           757
                                                                                                                                                                                                                    30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP1034785-A2
                                                                                                                                                18-NOV-1988;
                                                                                                                                                                                               8-NOV-1987;
                                           GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB18540;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB18540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            $\circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ{\circ}{\circ} \circ \ci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   It is the sequence encoded in the open reading frame of hepatitis C virus (HVV) cDNA inserts in clones K9-1 through 15e. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       635
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTHHADV1PVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GDSRGSLLSPRPISYLKGSSGGPLLCPACHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Caps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99.0%; Score 943; DB 10; Length 2436;
98.4%; Pred. No. 1.3e-88;
Live 3; Mismatches 0; indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus; clone 15e; clone k9-1; probe; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Purified hepatitis C virus - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Peptide encoded by composite hepatitis C cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 13; Figure 47-1 - 47-8; 139 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP90288 standard; protein; 2436 AA.
                                                                                                                                                                                                            870S-0122714.
870S-0139886.
880S-0161072.
880S-0191263.
880S-0263584.
880S-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                          Kuo G;
                                                                                                                                                                     88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (updated)
(updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 98.49
watches 179; Conscrvative
                                                                                                                                                                                                                                                                                                                                                                                                                                        Houghton M. Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1989-159274/22.
                                                                                                                                                                                                                                                                                                                                                                                         (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2436 AA;
                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAN92106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            useful in diagno
non-B hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RS 757
                                                                                                                                                                                                         18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
19-JUL-2001
01-NOV-1989
                                                                                                                                                                  18-NOV-1988;
                                                                                                                    31-MAY-1989
                                                                EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAP90288;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
```

DD

ò

qq

ò

g à

ò

ö

632

9

œ

```
121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                      This sequence represents the entire hepatitis C virus polyprotein. HCV is a member of the flavivirus family and appears to encode a basic polypeptide domain ("C") at the N-terminal of the viral polyprotein, followed by two glycoprotein domains ("E1", "E2/NS1"), upstream of the nonstructural genes NS2 through NS5. See also AAQ39134-48, AAR33982-4008 and AAR38088-89.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1026 LAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVTHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                            Immuno-reactive hepatitis C virus polypeptide compsns. - contg. at least 2 sequences from the first variable domain of distinct HCV isolates
gp53(BVDV)/gp55; hog cholera virus; pestivirus; NS1; flavivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Composite hepatitis C virus; HC-31/CDC/CHI; HCV; non-A non-B; synthetic antigens; blood screening.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                99.0%; Score 943; DB 14; 98.4%; Pred. No. 1.6e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                          (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Composite hepatitis C virus (HC-J1/CDC/CHI).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAR70230 standard; protein; 2894 AA.
                                                                                                                                                                                                                                                                                                             Disclosure; Fig 9; 106pp; English.
                                                                                                             92WO-US07683
                                                                                                                                        9105-0759575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 98.48
Matches 179; Conservative
                                                                                                                                                                                               Weiner AJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (updated)
                                                                                                                                                                                                                          WPI; 1993-117468/14.
                                                                                                                                                                   (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                      2816 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RS 182
                                                     W09306126-A1
                                                                                                           11-SEP-1992;
                                                                                                                                      13-SEP-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-MAR-2003
07-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EP644202-Al.
                                                                                                                                                                                               Houghton M,
                                                                                  01-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAR70230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR70230
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                        The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) antisense polynuclectide. The HCV is characterized by a positive stranded RNA genome which has 40% homology at the polypetide level to a HCV polyprotein. The antisense polynuclectide binds to cellular polynuclectides which enhance and/or are required for viral infectivity, replicative ability or chronicity. The antisense polynuclectides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity The composition also comprises an agent which causes viral RNA to be inactive. The composition is used for preventing HCV replication in a system. The present sequence is encoded by a novel HCV cDNA sequence, which is used in the course of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus; HCV; asymptomatic; chronically infected; cpitope; viral isolate; domain; immunological; cross-reactive; envelope protein; vaccine;
                                                                                                                                                                                                                         Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a sense strand of the virus genome, and selectively hybridises to it
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Match 99.0%; Score 943; DB 21; Length 2772; Local Similarity 98.4%; Pred. No. 1.6e-88; es 179; Conservative 3; Mismatches 0; Indels 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAR34009 standard; Protein; 2816 AA.
                                                                                                                                                                                                                                                                                Example; Fig 16; 75pp; English.
                                                    89US-0325338.
89US-0341334.
89US-0355002.
90EP-0302866.
                                                                                                                                                       Kuo G;
                          2000EP-0109602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (updated)
(first entry)
                                                                                                                                                       Choo 0,
                                                                                                                                                                                  WPI; 2000-566891/53.
                                                                                                                          (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2772 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV-1 polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ||
|RS || 1093
                                                                                                                                                                                              N-PSDB; AAA75296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RS 182
                          16-MAR-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
26-JUL-1993
                                                                   20-APR-1989;
18-MAY-1989;
                                                                                               16-MAR-1990;
                                                                                                                                                       Houghton M,
                                                     17-MAR-1989;
13-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR34009;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAR34009
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

1085

ö

Gaps

; 0

Length 2816; Indels

22-MAR-1995

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                           1 MAPITAYAQQTRGLLGCIITSIJGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 50
                                                                                                                                                                                                                                   AAR70230 is the composite hepatitis C virus (HC-J1/CDC/CHI) protein from which the synthetic HCV antigens described in AAR70210-R70229 were derived. These synthetic antigens can be used to screen blood, or blood products for the presence HCV, they can also be used in various specific assays for the detection of HCV antibodies, and antigens, or as immunogens.

(Updated on 25-MAR-2003 to correct PV field.)
                                                                                                                                                           Synthetic antigens for the detection of hepatitis C virus antibodies \cdot comprise portions of the HCV peptide sequence, for use in screening blood and blood products
                                                                                                                                                                                                                                                                                                                                                                                                   .
0
                                                                                                                                                                                                                                                                                                                                                                         99.0%; Score 943; DB 16; Length 2894; 98.4%; Pred. No. 1.7e-88;
                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                               Pollet D, Van Heuverswyn H;
                                                                                                                                                                                                                                                                                                                                                                                                 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Search completed: August 30, 2003, 19:12:19 Job time : 42.2251 secs
                                                                                                                                                                                                               Disclosure; Fig 1; 51pp; English.
                           94EP-0108611.
                                                  90EP-0124241.
                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 98.4'
Matches 179; Conservative
                                                                                      (INNO-) INNOGENETICS NV.
                                                                                                               Maertens G,
                                                                                                                                        WPI; 1995-116946/16.
                                                                                                                                                                                                                                                                                                                                                   2894 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RS 182
                                                  14-DEC-1990;
14-DEC-1990;
                           14-DEC-1990;
                                                                                                               Deleys RJ,
                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
Sense strand of the Hepstitis C virus Sense strand of HC Compiled HCV cont. Hepstitis C virus HCV polyprotein co HCV polyprotein co Hepstitis C virus Hepstitis C virus Hepstitis C virus Hepstitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                       Nucleotide sequence
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
CDNA encoding hepa
Hepatitis C virus
Nucleotide sequenc
Plasmid pr5His/HIV
HCV NS3 protease c
pnB183delta4A HT e
HCV Sclubilised NS
HCV solubilised NS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepalitis C virus
Hepalitis C virus
Hepatitis C virus
Hepatitis C virus
CDNA encoding hepa
                                                                                                 Hepatitis C virus
HCV-1 NS3/4a mutan
                                                                                                                                                        Combined open read
Hepatitis C virus
                                                                                                                                                                                      Combined open read
                                                                                                                                                                                                Combined open read
Composite hepatiti
Hepatitis C virus
                                                                                                                                                                                                                                         CDNA sequence comp
Nucleotide sequenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plasmid pET-BS(+)/
                                                                                                                                           DNA encoding HCV-1
                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS3 protease coding sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAT42394
AAT58401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX23259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABK87300
                                                                                                                                                                                                                                                                                                                                                                                                       AAX26737
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABK87286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4AT42389
                                                                                                                                                                                                                                                                                   AA01056
                                                                                                                                                                                                                                                                                                                                                                                            AAD31767
                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAV59361
                                                                                                                                                                                                                                                                                                AAA7529.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers 1..549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABK87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP.
                                                                                                 AAA70344 standard; DNA; 549
                                                                       DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                          %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                            2061
2061
2061
9185
9185
9502
9646
12980
16622
                                                                                                                                                                                                                                                                                                                                                                9401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                              19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                               \phi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA70344;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
                                                                                                                                                                                              Result
                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                **MODEL-frame+_poi.model -DEV-xip
-Q-/cgn2_1/USPTO_spool/US0996594/runat_29082003_151918_28302/app_query.fasta_1.2872
-Q-/cgn2_1/USPTO_spool/US09965594/runat_29082003_151918_28302/app_query.fasta_1.2872
-DB-N_Geneseq_19Jun03 -OFMT-1 -END--1 -MATRIX-blos.m62 -TRANS-human40.cdi
-LIST-45 -DOCALIGN-200 -THR_SCORE-pct -THR_MAX*.100 -THR_MAX*.00 -THR_MAX*.00 -THR_MAX*.00 -USPR-LOCAL.OUTFMT-pto -NORM-ext -HEAPSIZE-566 -MINLEN-0 -MAXIEN-200000000
-USRR-US09965594_eGCN_1_1.112.funat_29082003_151918_28302 -NCPU-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORE-0 -WAIT -DSPHLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_IIMEOUT-83 -THREADS-1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPEXT=0.5 -DELOP-6 -DELEXT=7
                                                                                          August 30, 2003, 19:13:57; Search time 169.009 Seconds (without alignments) 2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | SIDSI/qcgdata/geneseq.jennes.ncbl./Nal.980.DAT:*
| SIDSI/qcgdata/geneseq.jennes.ncbl./Nal.980.DAT:*
| SIDSI/qcgdata/geneseq.geneseqn.embl/Nal.990.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.992.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.992.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.993.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.993.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.994.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.995.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.995.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.995.DAT:*
| SIDSI/qcgdata/geneseq/geneseqn.embl/Nal.995.DAT:*
| SIDSI/qcgdata/geneseq/geneseqn.embl/Nal.999.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.999.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.999.DAT:*
| SIDSI/qcgdata/geneseqn.embl/Nal.000.DAT:*
| SIDSI/qcgdata/geneseq/geneseqn.embl/Nal.001B.DAT:*
                                                                                                                                                                              1 MAPITAYAQQTRGLLGCIIT......GVAKAVDFIPVESLETIMRS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N_Geneseq_19Jun03:*

1: /SIDS1/gcgdata/gcneseq/gcneseqn.embl/NA1980.DAT:*

2: /SIDS1/gcgdata/gcneseq/geneseqn.embl/NA1981.DAT:*

4: /SIDS1/gcgdata/geneseqn.embl/NA1982.DAT:*

5: /SIDS1/gcgdata/geneseqn.embl/NA1983.DAT:*

6: /SIDS1/gcgdata/geneseqn.embl/NA1984.DAT:*

7: /SIDS1/gcgdata/geneseqn.embl/NA1985.DAT:*

8: /SIDS1/gcgdata/geneseqn.embl/NA1985.DAT:*

9: /SIDS1/gcgdata/geneseqn.embl/NA1980.DAT:*

9: /SIDS1/gcgdata/geneseqn.embl/NA1987.DAT:*
                                                                                                                                                                                                                                                                                                                      5105512
             GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

    nucleic search, using frame_plus_p2n model

                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                           2552756 seqs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                       0.5
                                                                                                                                                                                                 BLOSUM62
Xgapop 10.0 , Xgapext (
Ygapop 10.0 , Ygapext (
Fgapop 6.0 , Fgapext )
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                US-09-965-594-1
953
                                                                                                                                                                                                                                                                                                                                                                                                                                             Command line parameters:
                                                                                                                                                   Title:
Perfect score:
                                                                                                                                                                                                          Scoring table:
                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Database :
                                                                                                                                                                              Sequence:
                                                                                                                                                                                                                                                                                           Searched:
                                                                                           Run on:
```

```
Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chien DY,
                                                                                                                                               541
                                                                                                                                                                                                             AAD29795;
121
                    361
                                         141
                                                             421
                                                                                                                                                                    RESULT 2
                                                                                                                                                                                         q
                                        ò
                                                            g
                                                                                  ò
                                                                                                    g
                                                                                                                           ò
                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80
                                                                                                                                                                                                                                                            The present sequence is the coding sequence for the Hepatitis C virus (HCV) NS3 protease enzyme. This protein is essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. NS4A is also needed for this process and inhibitors of the two proteins should act as antiviral treatments of HCV infection. This is useful'as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlySerScrAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                        549
0
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                           Sequence 549 BP; 83 A; 180 C; 142 G; 144 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                     Goldfarb
           protease"
                                                                                                                                    Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-1 (1-182) x AAA70344 (1-549)
                                                                                                                                                                                                                                            Example 1: Fig 9; 65pp; English.
                                                                                                                ပ္ပ
          /product= "NS3
                                                                                                                SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                       6.35e-80
953.00
100.00%
100.00%
                                                                                           99US-0115271.
                                                                                                                                    Weinheimer S,
                                                                       2000WO-US00345
                                                                                                                (BRIM ) BRISTOL-MYERS
                                                                                                                                                        WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                  P-PSDB; AAB15211
                               WO200040707-A1
                                                                                                                                     Wittekind M,
                                                                       06-JAN-2000;
                                                                                            08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                        complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
DB:
```

qq

ò Q çλ g ò a õ g

ó

```
420
                                                                                                160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present invention relates to hepatitis C virus (HCV) core antigen and NS (nonstructural) 3/4a antibody combination assay that can detect both HCV antigens and antiboddes present in a sample using a single solid matrix as well as immunoassay solid supports for use in the assay. The solid support is useful for detecting HCV infection in a biological sample. The present sequence is a DNA encoding HCV-1 NS)/4a mutant conformational antigen. This sequence is used in the exemplification
                                                                                                                                                                                            GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                          GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                     ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Coit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Immunoassay solid support useful for detecting hepatitis C virus infection in a biological sample, comprises at least one of HCV anti-core antibody and HCV NS3/4a epitope, bound to the support
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "HCV-1 NS3/4a conformational antigen"
"CDS does not include stop codon"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus; NS3/4a antigen; HCV infection; mutant;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              George-Nasciemento
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 2058 BP; 419 A; 634 C; 580 G; 425 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  antigen encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tandeske L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
1..686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               mutant conformational
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 2; Fig 4; 87pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                               BP
                                                                                                                                                                                                                                                                                                                                                                                                                                               AAD29795 standard; DNA; 2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-JUN-2000; 2000US-212082P. 02-APR-2001; 2001US-280811P. 02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               14-JUN-2001; 2001WO-US19369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /*tag= a
/product=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "C
/partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus type 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2002-179522/23.
P-PSDB; AAE18689.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                               181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                              CGTTCC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200196875-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV-1 NS3/4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20-DEC-2001
```

ä

3

```
The present invention relates to a new immunoassay solid support consisting essentially of at least one hepatitis C virus (ECV) NS3/4a conformational epitope and a multiple epitope fusion antiqen (MEPA), bound to the support. The NS3/4a conformational epitope and/or melection to the support. The NS3/4a conformational epitope and/or MEPA reacts specifically with anti-HCV antibodies present in a biological sample from an HCV-infected individual. The immunoassay of the invention is useful for detecting hepatitis C virus infection in a biological sample. The method of the invention provides a sensitive, accurate prevent transmission of HCV by blood and by blood products, or by prevent transmission of HCV by blood and by blood products, or by personal contact. Use of MS3/4a conformational epitope in combination with MEPA, provides a sensitive and reliable method for detecting early HCV seroconversion. Use of MEPA has the added advantages of decreasing masking problems, improving sensitivity in detecting antibodies by masking a greater number of epitopes on a unit surface area of substrate, and improving substrate Detection accuracy is increased and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            the incidence of false results is reduced because of the identification and the use of highly immunogenic HCV antigens which are present during the early stages of HCV seroconversion. The present nucleic acid sequence encodes the non-structural protein NS3/4a conformational epitope of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                         Immunoassay solid support, useful for detecting hepatitis C virus infection in biological sample, comprises HCV NS3/4a conformational epitope and multiple epitope fusion antigen bound to the support -
                                                                                                                                                             Coit
                                                                                                                                                             George-nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2058
180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 2058 BP; 419 A; 633 C; 581 G; 425 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-1 (1-182) x ABK15344 (1-2058)
                                                                                                                                                            Tandeske L,
                                                                                                                                                                                                                                                                                             Disclosure; Fig 3; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1.46e-78
                                                                             15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     946.00
100.00%
98.90%
99.27%
                                                     14-JUN-2001; 2001WO-US19156
                                                                                                                                                             Arcangel P,
                                                                                                                                                                                                   WPI; 2002-090228/12.
P-PSDB; AAU76377.
                                                                                                                                  (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                          Medina-selby A;
WO200196870-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                            20-DEC-2001
                                                                                                                                                             Chien DY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
                                                                                                                                                             9
                                                                                                                                                                                        40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion; immunoassay solid support; multiple epitope fusion antigen; MEFA;
                                                                                                                                  MetalaprollethralaTyralaGlnGlnThrargGlyLeuLeuGlyCysIleIlethr
                                                                                                                                                             ATGGCGCCCATCACGCCTACGCCCAGCAGAGAGGGGCCTCCTAGGGTGCATAATCACC
                                                                                                                                                                                      SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValScrThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /partial
/product- "HCV NS3/4a conformational epitope"
/note- "This sequence lacks a stop codon"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS3/4a conformational epitope gene seguence
            2058
1180
2
0
0
            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                 Gaps:
                                                                                                          (1-2058)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    non-structural protein; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BP
                                                                                                          US-09-965-594-1 (1-182) x AAD29795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABK15344 standard; DNA; 2058
            1.46e-78
946.00
100.00%
98.90%
99.27%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-MAY-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1..2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   111111
541 AGGTCC 546
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                      Best Local Similarity:
                                           Percent Similarity:
 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABK15344;
                                                                                                                                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                        21
                                                                                                                                                                                                                                                                                                                                                                                                    101
                                                                   Query Match:
DB:
               No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 3
                Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                    DP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                  ò
                                                                                                                                                          g
                                                                                                                                                                                        ò
                                                                                                                                                                                                           đ
                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

9

```
300
                                                       140
                                                                    420
                                                                                 160
                                                                                               480
                                                                                                             180
GlyAspSerArgClySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Coit,
                                                                                                                                                                                                                                   Immunoassay solid support; Hepatitis C Virus type-1; HCV-1; NS3/4a conformational epitope; multiple epitope fusion antigen; MEFA; anti-HCV antibody; NS3/4a conformational antigen; HCV infection; mutant; gene; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ú
                                                                                                                                                                                                                                                                                                                      /product- "NS3/4a conformational antigen"
/note- "This sequence lacks a stop codon"
                                                                                                                                                                                                                                                                                                                                                                                                                                                              George-Nascimento
                                                                                                                                                                                                                       encoding HCV-1 NS3/4a conformational antigen.
                                                                                                                                                                                                                                                                                           Location/Qualifiers
1..2058
/*tag= a
/partial
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŗ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Tandeske
                                                                                                                                                                                ВР
                                                                                                                                                                                DNA; 2058
                                                                                                                                                                                                                                                                                                                                                                                 15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                   ΰ
                                                                                                                                                                                                                                                                                                                                                                     2001US-0881654
                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                  GEORGE-NASCIMENTO
                                                                                                                                                                                                                                                                     Hepatitis C virus type 1. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                         COIT D. MEDINA-SELBY A.
                                                                                                                                                                                                                                                                                                                                                                                                             CHIEN DY.
ARCANGEL P.
TANDESKE L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2003-147573/14.
P-PSDB; ABG72261.
                                                                                                                                      181 ArgSer 182
                                                                                                                                                    AGGTCC 546
                                                                                                                                                                               ABX14410 standard;
                                                                                                                                                                                                                                                                                                                                           US2002146685-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                     14-JUN-2001;
                                                                                                                                                                                                           06-MAR-2003
                                                                                                                                                                                                                                                                                                                                                        10-0CT-2002
              241
                                                      121
                                                                   361
                                                                                 141
                                                                                               421
                                                                                                                                                    541
                                                                                                                                                                                              ABX14410;
                                                                                                                                                                                                                                                                                                                                                                                                             (CHIE/) (ARCA/) // (TAND/) 1 (GEOR/) C
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ďζ
                                                                                                                                                                                                                                                                                                                                                                                                                                         (COIT/) (MEDI/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Chien
                                                                                                                                                                ONA
                                                                                                                                   5 6
                                                                                                                      g
                                    Q
                                                                QQ
                                                                                 ò
                                                                                           οp
                                                                                                       5
           q
                         ò
                                                  ò
```

```
The present invention relates to immunoassays comprising Hepatitis C Virus (HCV) NS3/da conformational epitope and multiple epitope fusion antigen epitope and/or antigen (MEFA), bound to a solid support. The NS3/da epitope and/or the multiple epitope fusion antigen react with anti-HCV antibodies present in a biological sample from an HCV-infected individual. The immunoassays and methods of the invention are useful for detecting HCV infection in a biological sample. The inventive immunoassay solid support provides a sensitive and reliable method for detecting early HCV seroconversion. The assays can detect HCV infection caused by any St known genotypes of HCV. The use of the multiple epitope fusion proteins decreases masking problems, improves sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit area of substrate, and improves selectivity. The present sequence canceds HCV type 1 (HCV-1) NS3/da conformational antigen, a mutant of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80
 Hepatitis C Virus infection in C Virus conformational epitope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 AGCCTAACTGGCCGGGGACAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACTGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 GCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGGGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81 LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MethlaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ThrargGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet
                                                                                                                                                                                                                                                                                                                                  Sequence 2058 BP; 419 A; 633 C; 581 G; 425 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                            2058
180
2
0
0
                                                                                                                                                                                                                                                                                                                                                                                            Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                             Length:
Immunoassay solid support for detecting
biological samples, comprises Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-1 (1-182) x ABX14410 (1-2058)
                 biological samples, comprises Hepatit;
and multiple epitope fusion antigen
                                                              Disclosure; Fig 3A-3D; 45pp; English
                                                                                                                                                                                                                                                                                                    the HCV-1 NS3/4a polypeptide
                                                                                                                                                                                                                                                                                                                                                                            1.46e-78
946.00
100.00%
98.90%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161
                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                               Score:
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
```

AGGICC

a

S

```
AAN90327
IID AAN9
IID AAN9
XX
AX
XX
XX
DT 25-19
DI 11-1-1
DIE HEPE
XX
XX
XX
HEPE
XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
8 3 8 8 8 8 8 8
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 14i. 11b, 7f, 7e, 8h 33c, 40b, 35b, 35b, 35b, 35b, 10c, 10c, 8f, 31f, 31g and bepatitis C 37b, 35b, 35b, 35b, 10c, 10c, 31c, 31g, 31d, 31g, and 9c of hepatitis C virus (HCV) cDNA. In creating the composite sequence the following heterogeneities were considered. Clone 33c contains a sequence of 80c base pairs which overlaps the clones 31c choices, 10c and 37c. In clone 33c, as well as in 5 other coverlapping clones, nucleotide $789 is a G. However, in clone 37b the corresponding nucleotide is an A. This heterogenity may have important ramifications for protein folding. Nucleotide $2 in clone 37b the corresponding nucleotide is an A. This clone 38h is a T which may represent a cloning artifact because the corresponding residue in clone 7e and in 3 other overlapping clones is an A. The 3-terminal nucleotide in clone 8f is represented as a T than a G because the corresponding residue in clone 31f and in 2 other overlapping clones is a T. The 3' terminal sequence of clone 33f is represented as ATTC, as is found in the corresponding sequence in clone 35g and in 2 other overlapping clones, rather than as TIGC, as 15 found in clone 31f and an A. The therefore the corresponding sequence in clone 35g and in 2 other overlapping clones, rather than as TIGC, as 15 found in clone 31f and a ATTC.
                                                                                                                                                      Combined open reading frames of the hepatitis C virus (HCV) cDNA in clones 141 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         clones is an A. The 3' terminus of clone 14 is depicted as TA rather than AA because the corresponding dinucleotide in clone 11b and 3 other clones is TA. Potential cloning artifacts have been omitted and instead the corresponding sequences in non-5'-terminal regions of multiple overlapping clones are shown. AANO2097 could be used as a source of oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             because the corresponding residue in clone 33f and 2 other overlapping
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 3; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                         Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                       Location/Qualifiers
3..5300
/*tag= a
                                              AAN92097 standard; DNA; 5300 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      87US-0122714
                                                                                                                                                                                                                                                                                                                                                                                                                       8BEP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88US-0271450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Purified hepatitis C virus
                                                                                                                             (first entry)
                                                                                                               (nbdated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1989-159274/22.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     P-PSDB; AAP92041
                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                      18-NOV-1988;
                                                                                                             25-MAR-2003
                                                                                                                             02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                         31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                        EP318216-A
                                                                                AAN92097;
                                AAN92097
```

```
1229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1289
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1290 GGTGATAGCAGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCCTCG 1349
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyTlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161 ThrargGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
                                                                                                                                                                                                                                                                                                                                                                                                          40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
immuno- assay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B \,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             990 AGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCTAACTGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81 LysaspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 GlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLySGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                          SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg
                                                                                                                                                                                                                                                                                                                                      1 MetAlaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                           Sequence 5300 BP; 1047 A; 1606 C; 1515 G; 1130 T; 2 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus; composite cDNA; probe; vaccine.
                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                    (Updated on 25-MAR-2003 to correct PR field. (Updated on 25-MAR-2003 to correct PI field.
                                                                                                                                                                                                                                          Indels:
                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                 US-09-965-594-1 (1-182) x AAN92097 (1-5300)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus composite probe.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAN90327 standard; cDNA; 5360 BP.
                                                                                                                                                                  8.9e-78
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1470 AGGTCC 1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 ArgSer 182
                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25-MAR-2003
                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAN90327;
                                                                                                                                                                                                                                                                                                                                                                                                            21
                                                                                                                                                                                                                                              Query Match:
                                                                                                                                                                      ..
Q
```

Key

```
It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 12f to 15e of hepaticis C virus (HCV) CDNA. It can be used to make oligomeric DNA hybridisation probes t detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                  Combined open reading frames of the hepatitis C virus (HCV) CDNAs from
          141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                  Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                          non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 3; Figure 32-1 - 32-7; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
3..6905
/*tag= a
                                                                                                                                                                                                                                                   BP.
                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; HCV; non-A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                  AAN92103 standard; DNA; 6905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                          (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                 clones 12f through 15e.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Choo QL,
                                                                                                                                                                                    1470 AGGICC 1475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1989-159274/22.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (CHIR ) CHIRON CORP.
                                                                                                                                                                  ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Houghton M,
                                                                                                                                                                                                                                                                                                         25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EP318216-A.
                         1290
                                                                                 1350
                                                                                                                                                                  181
                                                                                                                                                                                                                                                                              AAN92103;
121
                                                                                                                                                                                                                                                                                                                                                                                                                                               Rey
                                                                                                                                                                                                                         RESULT 7
                                                                                                                                                                                                                                                                 d
                                                                              a
                                                                                                             ô
                                                                                                                                                                  ö
                       a
                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :::[[[[[[[[]]]]]]]]]]
1170 CAAGACCTTGTGGGCTGGCCGCTCGCAAGGTAGCCGTCATTGACACCCTGCACTTGC 1229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                96
                                                                                                                                                                                                                                                                                                                                                                          The sequence shows the composite cDNA sequence derived from the aligned hepatitis C virus (HCV) cDNA's in clones 14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33f and 39c. The cDNA encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The cDNA can be used to design probes, or to synthesise polypeptides, which are used to dispasse HCV induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP90158, AAN90303-26, and AAN90328-36.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lysaspi.euValGiyTrpProAlaProGlnGlySerArgSer..euThrProCysThrCys
                                                                                                                                                                                                                                                                                  Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 5360 BP; 1060 A; 1622 C; 1532 G; 1145 T; 1 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5360
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conscrvative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-1 (1-182) x AAN90327 (1-5360)
                                                                                                                                                                                                                                                                                                                                                  Disclosure; Fig. 26; 174pp; English.
Location/Qualifiers
                                                                                                                                                                                                                                      ö
                                                                                                                                     87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9.03e-78
943.00
100.00%
98.35%
98.95%
                                                                                                             88GB-0027024
                                                                                                                                                                                                                                        Kuo
                                                                                                                                                                                                            (CHIR ) CHIRON CORPORATION
                            d
          3..5360
/*tag-
                                                                                                                                                                                                                                     Choo QL,
                                                                                                                                                                                                                                                               WPI; 1989-215054/30.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                   30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                     Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                             18-NOV-1988;
                                                     GB2212511-A.
                                                                                 26-JUL-1989
                                                                                                                                        18-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Match:
```

Query

ŏ a ò a ò a ò q ò ď ò

ಭ

```
CORP
                                                                                  WPI; 1989-159274/22
                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                         P-PSDB; AAP92050
                                                          (CHIR ) CHIRON
                18-NOV-1987;
30-DEC-1987;
26-PEB-1988;
06-MAY-1988;
                                                                                                                                                                                                         Alignment Scores:
                                              14-NOV-1988;
     8-NOV-1988
                                         26-OCT-1988
                                                                                                                                                                                                                                                                                                                                                                                1968
                                                                                                                                                                                                                                                                                                                                                                                                        2028
                                                                                                                                                                                                                                                                                             21
                                                                                                                                                                                                                                                                                                                                                                                            101
                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                 ..
⊗
                                                                                                                                                                                                                                                                     င်
                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                              DP
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                a
1322
                                                                                                                                 1683 ACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGACCTAGAGACAACCATG 1742
                                                         121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                         161 ThrangGlyValalaLysAlaValAspPhe1leProValGluSerLeuGluThrThrMet 180
                                                                                  04
                                                                                                                                                                                                                                                                                                                                                             open reading frames of the hepatitis C virus (HCV) cDNAs from
                                                                                         21 SerbeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                hepatitis; NANBH
   6905
179
3
0
0
    Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                               x AAN92103 (1-6905)
                                                                                                                                                                                                                                                                                                                                                                                 non-B
                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers 3..7310
                                                                                                                                                                                                                                                                                                                                                                                 HCV; non-A,
                                                                                                                                                                                                                                                                                                                    AAN92106 standard: DNA; 7310
     23e-77
           943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                    clones K9-1 through 15e
                                                                                                                                                                                                                                                                                             1743 AGGTCC 1748
                                                                                                                                                                                                                                                                                 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus;
                                               US-09-965-594-1 (1-182)
                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
               Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                            25-MAR-2003
02-MAR-1990
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                             31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                 EP318216-A
                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                 AAN92106;
                                                                                                                                                                                                                                                                                                                                                               Combined
      ..
9
                                                                                                                                                                                                                                                                                                                                                                                                         Key
                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                           g
                                                                                                                                            qq
                                                                                                                                                                                           g
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                 ð
                                                                     q
                                                                                                                                                                   đ
                                                                                                                                                                                 ò
                                                                                                                                                                                                       à
                                                                                                                                                                                                                                 ő
                                                          ò
                                                                                  ò
                                                                                           g
                                                                                                         οy
                                                                                                                   q
                                                                                                                              ò
                                                                                                                                                          5
```

```
1847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2027
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1967
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGly1lePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0
                                                                                                                                                                                                                                                                                                                                                                             It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus (HCV) cONA. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1788 AGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTGTAACTGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81 LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; 2058 G; 1540 T; 4 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7310
179
3
0
0
                                                                                                                                                                                                                                                                                                            and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                            claim 3; Figure 47-1 - 47-8; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-1 (1-182) x AAN92106 (1-7310)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 7310 BP; 1491 A; 2217
                              87US-0122714.
87US-013986.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                 Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.32e-77
943.00
100.00%
98.35%
98.95%
88EP-0310922
                                                                                                                                                                                                                                                                                         Purified hepatitis C virus
                                                                                                                                                                                               Houghton M, Choo QL,
```

g

qq ò

õ

```
1968 CAAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACTTGC 2027
               141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                         MetalaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr
                                                                                                                                                       61 GlyThrArgThr1leAlaSerProLysGlyProVal1leGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                              121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                            41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine; antibodies; immunoprophylaxis; sera; serum; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               standard; cDNA to mRNA; 7310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus clone genome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Reyes GR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91US-0681703.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         900S-0505611.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GENE-) GENELABS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2268 AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Moeckli R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1995-302120/39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C5-APR-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-APR-1990;
09-OCT-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-MAR-2003
15-AUG-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US5443965-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22-AUG-1995
                                                                                                                                                                                                                                                                                                                                                                                        2148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAQ98221;
 -
                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAQ98221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kim JP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    aa098221
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                        a
                                                                                                                                                       à
                                                                                                                                                                               g
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                   a
                    8
                                               9
                                                                         q
                                                                                                     8
                                                    ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The sequence shows a composite hepatitis C virus (HCV) cDNA, derived be aligning clones k9-1 through 15e in 5'-3' direction. The CDNA encodes antiquens which react with antibodies in patients with non-6 non-8 hepatitis (NANBH). The CDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH.
Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                      161 ThrArgGlyValAlaLysAlaValAspPhcIleProValGluSerLeuGluThrThrMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP90288, and AAM90303-35.

(N.B. This record was resubmitted to correct crrors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                  ds.
                                                                                                                                                                                                                                                                                                Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7310
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                         Composite hepatitis C virus (HCV) CDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; fig 47; 235pp; English.
                                                                                                                                                                    AAN90336 standard; DNA; 7310 BP
                                                                                                                                                                                                                                                                                                                                                                                                                            87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1,32e-77
943.00
100.008
98.358
98.958
                                                                                                                                                                                                                     (updated)
(updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                       88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Houghton M, Choo QL,
                                                                                                                 AGGICC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1989-215054/30.
P-PSDB; AAP90288.
                                                                                          Argser 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                          Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                         30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                      18-NOV-1988;
                                                                                                                                                                                                                                 19-JUL-2001
01-NOV-1989
                                                                                                                                                                                                                      25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                   3B2212511-A.
                                                                                                                                                                                                                                                                                                                                                                            26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-1987
                                                                                        181
                                                                                                                   2268
                                                                                                                                                                                              4AN90336;
```

2087

140

120

1967

80

100

2207

to of

New nucleic acids encoding hepatitis C virus antigens - used develop prods. for detection of HCV-infected sera and prodn.

(1-7310)

US-09-965-594-1 (1-182) x AAN90336

Score

```
CDNA sequence compiled Hepatitis C virus CDNA clones
                                                                                                                                                                                                            CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                              Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                               17-MAR-1989;
20-APR-1989;
                                                                                                           EP1034785-A2
                                                                                                                                                                                          16-MAR-1990;
                                                                                                                                                                                  18-MAY-1989;
                                                                                                                            13-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ...
...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score:
සු
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2027
                                                                                                                                                                                                                                                                                                                                                                                                                     2088 GGTGATAGCAGGGGCAGCCTGCTGTCGCCCGGCCCATTTCCTACTTGAAAGGCTCCTCG 2147
                                                                                                                                                                                                                                                                                                                                                                                                                                               7.60
                                                                                                                                                                                                                                                                                                                                                                                                            GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
                                                                                                                                                                                                                                                                                                                                                                         LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys 100
                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                   09
                                                                                                                                                                                                                                                                                                                                                                                  1968 CAAGACCTTGTGGGCTGGCCCCCTCCGCAAGGTAGCCGCTCATTGACACCTGCACTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGly11ePhcArgAlaAlaValCys
                                                                                                                                                                                                                                                                 SerLeuThrGlyArgAspi.ysAsnGlnValGluClyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                   41 AlaGinThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                   Hepatitis C virus (HCV) antiqens can be used for detecting HCV infected sera and individuals infected with HCV. They can also be used in an anti-HCV vaccine or for the production of anti-HCV antibodies which can be used for passive immunoprophylaxis. The antigens consistently identify more HCV positive serum samples with a high degree of specificity. See AAQ98202-14 and AAR81939-51. (Updated on 25-MAR-2003 to correct PF field.)
                                                                                                                   Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 other;
                                                                                                                                             7310
179
3
0
0
                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                 Indels:
Gaps:
                                                                                                                                                                                                             US-09-965-594-1 (1-182) x AAQ98221 (1-7310)
                    Example 4; Figure 11; 71pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ВЪ
vaccines and anti-HCV antibodies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CDNA; 8316
                                                                                                                                             1.32e-77
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JAN-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2268 AGGTCC 2273
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                standard;
                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                      Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                               1728
                                                                                                                                                                                                                                                                                  1788
                                                                                                                                                                                                                                                                                                                                                       1908
                                                                                                                                                                                                                                                                                                                                                                                                            101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA75296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA75296;
                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA75296

1D AAA7

XX

AC AAA7

XX

DT 15-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
 ò
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

```
2734 CIGGCGCCCAICACGCCGIACGCCCAGCAGACAAGGGGCCICCIAGGGTGCATAAICACC 2793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) antisense polynuciacutde. The HCV is characterized by a positive stranded RNA genome which has 40% homology at the polypeptide level to a HCV polyprotein. The antisense polynuciacuted binds to cellular polynuciacutise which enhance and/or are required for viral infectivity, replicative belinity or chronicity. The antisense polynuciacides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity. The composition also comprises an agent which causes viral RNA to be inactive. The composition is sed for preventing HCV replication in a system. The present course of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag= a
/note= "partial sequence; no termination codon given"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a sense strand of the virus genome, and selectively hybridises to it -
Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
viral infectivity; viral replication; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; 2345 G; 1771 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8316
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-1 (1-182) x AAA75296 (1-8316)
                                                                                                                          Location/Qualifiers
1..8316
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 8316 BP; 1671 A; 2529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example; Fig 16; 75pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1.55e-77
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                 89US-0325338.
89US-0341334.
89US-0355002.
90EP-0302866.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                       16-MAR-1990; 2000EP-0109602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Houghton M, Choo Q,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-566891/53.
P-PSDB; AAB18540.
```

us-09-965-594-1.rng

```
Score
                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                                                                                   6
                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                         gg
                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                              \ddot{o}
                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
ò
                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                     ò
                                                                                                     /tansl_except. (pos:1588..1589; aa:Leu)
/note= 'this codon has an apparent 1 nucleotide deletion,
which alters the reading frame'
/transl_except= (pos:1647..1650; aa:Pro)
/note= 'this codon has an apparent 1 nucleotide
insertion, which alters the reading frame: insertion, which alters the reading frame: this
present in the formal sequence
present in the formal sequence listing of the
specification.
                                                      101 GlyscrSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
                                                                                                                                                                                                                                                Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH; HCV infection; vaccine; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Kolberg JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Irvine BD,
                                                                                                                                                                                                                                                                                Location/Qualifiers
268..9132
                                                                                                                                                                                                                                                                                                                                                                                                                                               Houghton M, Irvilito I, Weiner AJ;
                                                                                                                                                                                                                                      Nucleotide sequence of HCV-1 ORF.
                                                                                                                                                                                           AAZ07656 standard; DNA; 9133 BP
                                                                                                                                                                                                                                                                                                                                                                                                  89US-0408045.
89US-0456142.
90EP-0310149.
                                                                                                                                                                                                                                                                                                                                                                                       99FP-0101746
                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                             /*tag- b
                                                                                                                                                                                                                    (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                          Han J, ... Saito J
                                                                                                                                                                  3274 AGGTCC 3279
                                                                                                                                                                                                                                                                                                                                                                                                                           (CHIR ) CHIRON CORP. (OYAA/) OYA A.
                                                                                                                                                      181 ArgSer 182
                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                  15-SEP-1989;
21-DEC-1989;
17-SEP-1990;
                                                                                                                                                                                                                  20-MAR-2003
08-NOV-1999
                                                                                                                                                                                                                                                                                                                                                                                       17-SEP-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Miyamura T,
                                                                                                                                                                                                                                                                                                                                                                           01-SEP-1999
                                                                                                                                                                                                                                                                                                                                                              EP939128-A2
                                                                                                                                                                                                        AAZ07656;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Cha T,
                                                                                                                                                                                    g
                                                                                                                          ò
                                                                                                                                       QQ
                                                                                                                                                     ŏ
                                                                                                                                                               q
                 a
                             õ
                                       Q
                                                     õ
                                                               g
                                                                              õ
                                                                                       q
                                                                                                  ò
```

```
3582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3642
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100
                                                                                                                                                                     and J. These two isolates comprise nuclectide and anino acid sequences that are distinct from the HCV isolate HCV-1. The nucleotide sequences that are distinct from the HCV isolate HCV-1. The nucleotide sequences may be used to detect non-A, non-B HCV (NAMBH) polynucleotides by screen blood donors, donated blood and blood products for this infection. The isolates may also be used to isolate other naturally occurring variants of the virus. The polypeptides may be used as a vaccine for administration to patients to protect against infection with NAMBH. The (Updated on 20-MAR-2003 to correct PF field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3583 CARGACCTTGTGGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCTGCACTTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MetalaProlleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIlelleThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81 LysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161 ThrargGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet
                                                                                                                                                           The invention provides two new isolates of hepatitis C virus (HCV), J1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3463 GCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGGGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                           C Virus isolates, useful for diagnosis of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9133
179
3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-1 (1-182) x AAZ07656 (1-9133)
                                                                                infections and development of vaccines
                                                                                                                    Disclosure; Fig 12; 132pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.74e-77
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AGGICC 3888
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 ArgSer 182
1999-480843/41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
  WPI; 1999-480843,
P-PSDB; AAY14975
                                                             New Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3883
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
```

us-09-965-594-1.rng

0

AAQ05956;

RESULT 13 AAQ05956

misc_RNA

misc_RNA

3514

40

3574

9

```
121 GlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
                                                                                   3395 CTGGCGCCCATCACGGCGTACGCCCAGCAGACAGGGGCCTCCTAGGGTGCATAATCACC 3454
                                                                                                                                                                                    3755 GGTGATAGCAGGGAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCCTCG
                                                                                                                             41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla
                                                      SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
Gaps:
                           US-09-965-594-1 (1-182) x AAQ05956 (1-9185)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAQ10566 standard; DNA; 9185 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus strain 1 DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            90EP-0309120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       890S-0398667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (updated)
(first entry)
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             viral infections; ss
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3935 AGGTCC 3940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 ArgSer 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 1991-059670/09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-AUG-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-AUG-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-MAR-2003
29-APR-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27-FEB-1991.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3P414475-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Weiner AJ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAQ10566;
                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAQ10566
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                     οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                   a
                                                                                                                                                                      ò
                                                                                                                                                                                                a
                                                                                                                                                                                                                             δ
                                                        à
                                                                                                                ò
                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV CDNA libraries were constructed using pooled serum from a chimpanzee with chronic HCV infection. A lambda 9111 library was screened with probes derived from previously isolated clones. The ORF is derived from the overlapping clones bild, ag30a, CA205a, 12f, 14i, 11b, 7f, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 2cc, 14c, 8f, 31f, 39c, 35f, 19g, 26g, 15e, b5a and 16jh. These clones extend the sequence of the HCV genome reported in EP-318216. The upstream region from nucleotides -319 to -1348 (-1.1667 in this file) is covered by clones bilda, 18g, ag30a, CA205a, CA205a, CA205a, pida, cA167b, CA156e, CA84a and CA59a; nucleotides 8659-8866 (-8978-9185 in this file) are covered by clones b5a and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus DNA - used for producing probes, polypeptide(s), antibodies and anti-sense polynucleotide(s) for diagnosis and therapy.
                                                                                                                                                                                                                                                                                                     within this region is claimed"
                                                                                                                                            cDNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 other;
                                                                                                                                                                                                                                                                                                                                             /note="encodes an epitope that is claimed"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                            Sense strand of the compiled Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               See also AAQ05955.
(Updated on 25-MAR-2003 to correct PA field.)
                                                                                                                                                                        Hepatitis C virus (HCV); antiviral agent; ss
                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Fig 17; 83pp; English.
                                                                                                                                                                                                                                                        /*tag- a
1.1667
/*tag= b
/note="epitope w
8978.9185
                                        AAQ05956 standard; DNA; 9185 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           89US-0355002.
89US-0325338.
89US-0341334.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.75e-77
943.00
100.00%
98.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                  90EP-0302866
                                                                                                   (updated)
(first entry)
                                                                                                                                                                                                                                            320..9185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1990-284418/38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
Query Match:
                                                                                                                                                                                                   virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P-PSDB; AAR08124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-MAY-1989;
17-MAR-1989;
20-APR-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                  16-MAR-1990;
                                                                                                  25-MAR-2003
23-JAN-1991
                                                                                                                                                                                                   Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                      19-SEP-1990
                                                                                                                                                                                                                                                                                                                                                                             EP388232-A
```

```
(first entry)
                                                                                                                                                                                                                                               Houghton M, Choo Q,
                                                                                                                                                                                                                                                                 WPI; 2000-566891/53.
                                                                                                                                                                                                                              (CHIR ) CHIRON CORP
                                          Hepatitis C virus;
viral infectivity;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity:
Query Match:
                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                          P-PSDB; AAB18541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                            20-APR-1989;
18-MAY-1989;
16-MAR-1990;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                               EP1034785-A2
        15-JAN-2001
                                                                                                                                                 13-SEP-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                    3874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3395 CTGGCGCCCAICACGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAAICACC 3454
                                                                                                                                                                                                                                                                                                                                                                                                                                                             140
                                                                                                                                                                                                                                                                                                                                                                                                                            GlySerSerAspLeuTyrLeuValThrArqHisAlaAspValIleProValArgArgArg 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
                                                                                                                                                                                                                                                                                                                                                       80
                                                                                                                                                                                                                                               1 MetalaProlleThralaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
                                                                                                                                                                                                                                                                                                                                                       GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyAspScrArgGlySerLcuLeuSerProArgProlleScrTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This is a hepatitis C virus (HCV) composite chWA sequence, deduced using overlapping clones. a compsn. contg. the artigonic protein bodies (Abs) and for screening an agent which inhibits HCV articles. A cell line infected with this virus can be used as a action. A cell line infected with this virus can be used as a for artigons. The antigon is useful for preparing vaccines for treating viral infections. See also AAQ10557.
      Cell lines infected with hepatitis C virus - are used as source of antigens for detection of HCV antibodies, for vaccines, and for screening anti-viral agents
                                                                                                                                        Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 other;
                                                                                                                                                                  9185
179
3
0
0
                                                                                                                                                                 Length:
Matches:
Conscrvative:
Mismatches:
Indels:
                                                                                                                                                                                                                              US-09-965-594-1 (1-182) x AAQ10566 (1-9185)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA75297 standard; cDNA; 9185 BP.
                                                                                                                                                               1.75e-77
943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3935 AGGTCC 3940
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ArgSer 182
                                          fig l;
                                                                                                                                                                                            Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                   Percent Similarity
                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                        3575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                            101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA75297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
AAA75297
ID AAA75:
XX
AC AAA75:
                                                                                                                                                                   No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
ò
                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
1 MetalaProlleThralaTyralaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
3395 CTGGCGCCCATCACGCGTACGCCCAGCACAAGGGCCTCCTAGGGTGCATAATCACC 3454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) antisense polynucleotide. The HCV is characterized by a positive stranded RNA genome which has antisense polynucleotide by a positive stranded RNA genome which has antisense polynucleotide binds to cellular polynucleotides which enhance and/or are required for viral infectivity, replicative ability or chronicity. The antisense polynucleotides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity. The composition also comprises an agent which causes viral RNA to be inactive. The composition is used for preventing HCV replication in a system. The present sequence represents a novel HCV cDNA sequence, which is used in the
                                                                                                                                                                                                                                                      /*tag= a
/note= "partial sequence; no termination codon glven"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sense
                                                   HCV, antisense polynucleotide; polyprotein; viral replication; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a strand of the virus genome, and selectively hybridises to it
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 other
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9185
179
.3
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches:
Conservative:
Mismatches:
Sense strand of HCV encoding a polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-1 (1-182) x AAA75297 (1-9185)
                                                                                                                                                                                                Location/Qualifiers
320..9184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Example; Fig 17; 75pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89US-0341334.
89US-0355002.
90EP-0302866.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89US-0325338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        943.00
100.00%
98.35%
98.95%
                                                                                                                                                                                                                                                                                                                                                                                                                                                         16-MAR-1990; 2000EP-0109602
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      course of the invention.
```

```
3935 AGGTCC 3940
                        181 ArgSer 182
   61
     3575
   S S
                        õ
 g
ŝ
```

Search completed: August 30, 2003, 19:47:38 Job time: 188.009 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 9.01384 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-1 953 1 MAPITAYAQQIRGLIGCIIT......GVAKAVDFIPVESI.HTTMRS 182

Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 6% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

No. Score Match Length DB ID Description			æ			SUMMARIES	
200	Result	6	Query	4	2	£	
1 943 99.0 3011 1 POLG_HCVH P26664 2 97.3 3011 1 POLG_HCVH P22866 3 93.6 3010 1 POLG_HCVH P2666 891 93.5 3010 1 POLG_HCVT P2666 884 93.1 3010 1 POLG_HCVT P2666 714 74.9 3033 1 POLG_HCVT P2666 84 98.8 3010 1 POLG_HCVT P2666 84 98.8 1000 1 POLG_HCVT P2666 84 485 1 HOA_HART P2666 93.3 1 POLG_HCVT P2666 94 83.1 1 HPOR_ARRTH P2666 95 84 437 1 DEGI_RATH P2666 74.5 78 415 1	NO.	arose	Match	Length	9	01	Description
2 97.3 3011 1 POLG_HCVTW PO266846 8 93.6 3010 1 POLG_HCVTW PO26662 8 93.6 3010 1 POLG_HCVTW PO26662 8 93.1 3010 1 POLG_HCVTW PO26662 8 93.1 3010 1 POLG_HCVTW PO26662 8 94.8 3033 1 POLG_HCVTW PO26660 9 1.2 74.7 3033 1 POLG_HCVTW PO26660 9 1.2 74.7 3033 1 POLG_HCVTW PO26660 9 1.2 74.7 3033 1 POLG_HCVTW PO26660 9 1.2 1.2 1.2 1.2 1.2 9 1.2 1.2 1.2 1.2 1.2 9 2.5 1.2 1.2 1.2 1.2 1 1.2 1.2 1.2 1.2 1.2 1.2 1 1.2 1.2 1.2 1.2 1.2 1.2	н	4		3011	7	POLG_HCV1	P26664 h genome po
892 93.6 3010 1 POLG HOVTW P26663 887 93.5 3010 1 POLG HOVTW P26663 887 93.5 3010 1 POLG HOVJA P26663 884 92.8 3010 1 POLG HOVJA P26663 884 92.8 3010 1 POLG HOVJA P26661 884 92.8 3033 1 POLG HOVJA P26661 887 9.1 209 1 PAAD PSEAE P2661 882 8.6 452 1 AAMP HUMAN P267 883 8.6 485 1 V136 TREPA P2681 884 885 1 V136 TREPA P2681 885 8.6 485 1 V136 TREPA P2681 886 8.8 485 1 V136 TREPA P2681 887 9.1 209 1 PAAD PSEAE P2681 888 8.6 485 1 V136 TREPA P2681 889 8.6 485 1 V136 TREPA P2681 880 8.8 485 1 V136 TREPA P2681 880 8.8 485 1 V136 TREPA P2681 881 8.6 485 1 V136 TREPA P2681 882 8.6 485 1 V186 883 8.6 485 1 V186 884 8.5 1 V186 885 8.6 485 1 V186 886 8.6 485 1 V186 887 9.1 2 V18 888 8.6 485 1 V186 889 8.6 485 1 V186 891 1 V11 WAN WOUSE 891 1 V11 WAN WOUSE 892 8.6 627 1 CALH WOUSE 893 1 VPRT SMRYH P2681 893 1 VPRT SMRYH P2681 893 1 VPRT SMRYH P2681 893 1 VRT SMRYH P2681 894 V17 895 V17 895 V17 896 V186 V186 V186 896 V17 897 V186 V186 V186 898 V186 V186 V186 V186 898 V186 V186 V186 V186 V186 V186 V186 V18	7	927		3011	-	POLG_HCVH	ב
4 891 93.5 3010 1 POLG_HCVJA P2665 5 884 92.8 3010 1 POLG_HCVJA P2666 7 74 74.9 3033 1 POLG_HCVJG P2666 84 92.8 3010 1 POLG_HCVJG P2666 84 92.8 3010 1 POLG_HCVJG P2666 84 92.8 3010 1 POLG_HCVJG P2666 84 9.8 321 1 HOA_ARATH P2666 82 8.6 485 1 AAB_HABAN Q1886 80.5 8.4 437 1 DEGI_ARATH Q1886 80.5 8.4 437 1 DEGI_ARATH P4689 80.5 8.4 415 1 PAB_HABAN Q1805 80.5 8.4 415 1 PAB_	m	892		3010	_	POLG_HCVTW	h genome
5 887 93.1 3010 1 POLG_HCVJT PDGG_GCS 6 884 92.8 3010 1 POLG_HCVJT PDGGGGS 7 74 74.9 3033 1 POLG_HCVJT PDGGGGS 8 8.1 4.5 1 PAD_PSEAE PDGGGGS 9 1.2 7.7 3033 1 POLG_HCVJT PDGGGGS 9 1.2 20.9 1 PAD_PSEAE PDGGGGG PDGGGGG 9 1.2 1.4 4.3 1 HNOA_ARATH QDGGG QDGGG 1 8.6 4.5 1 AAMP_HUAAN QDGGGG QDGGGG QDGGGG QDGGGG 4 4.3 1 DEGL_ARATH PDGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	∢	891		3010	7	POLG_HCVBK	h genome
6 884 92.8 3010 I POLG_HCVJT 002269 712 74.7 3033 I POLG_HCVJE P26661 9 71.2 74.7 3033 I POLG_HCVJE P26661 9 7.1 3033 I POLG_HCVJE P26661 9 8.6 452 I AAMP_HDAA P26661 9 8.6 485 I X36 TREPA Q958217 9 8.6 485 I X36 TREPA Q91172 4 3.5 1 POLG_HAATH Q958217 Q958217 4 3.5 1 POLG_HAATH Q05869 5 7.9 253 1 CAC3_BOVIN P05805 6 7.7 7.8 911 1 PRIL_BABIT P05805 7 7.8 911 1 PRIL_BABIT P05805 7 7.8 911 1 PRIL_BABIT P05905 7 7.9 257 1 RAM_HOMAN P05906 7 7.6 157 <t< td=""><td>ស</td><td>887</td><td></td><td>3010</td><td>~4</td><td>POLG_HCVJA</td><td>h genome</td></t<>	ស	887		3010	~4	POLG_HCVJA	h genome
7 714 74.9 3033 1 POLG_HCVJ6 P2666 8 8 8.8 323 1 POLG_HCVJ6 P2666 8 9 9.1 2033 1 POLG_HCVJ6 P2666 8 9 9.1 2033 1 POLG_HCVJ6 P2666 8 9 9.1 203 1 POLG_HCVJ6 P2666 8 9 9.1 203 1 PADA_ANXTH P23 8 0.5 8.4 437 1 DEGI_NRATH P23 8 0.5 8.4 437 1 DEGI_NRATH P26805 7 7.9 253 1 CAC3_BOVIN P05805 7 4.5 7.8 415 1 LP3_RABIT P46833 7 4.5 7.8 776 1 HYPF_AZOVI P46893 7 4.5 7.8 911 1 PADI_NETMB	9	884		3010	-4	POLG_HCVJT	h denome
8 712 74.7 3033 1 POLG, HOVJ6 PP6660 9 84 9.1 209 1 HOAD_PSEAE O95617 1 82 8.6 452 1 AAMP_HOMAN O11665 2 8.6 452 1 AAMP_HOMAN O11665 2 8.6 452 1 AAMP_HOMAN O11665 3 8.6 452 1 AAMP_HOMAN O11665 4 475 1 DEGL_ARATH O22609 5 7.9 253 1 CAC3_BOVIN PP68805 6 7.7 7.8 7.6 1 HYPP_AZOVI PP68805 7 7.8 9.1 1 TBILLINITH PP6893 PP68905 8 7.7 263 1 RPALE_MITICO O99075 O99075 8 7.7 263 1 RPALE_MITICO O99075 O980710 1 7.5 7.6 627 1 GRAM_HOMAN PP5104 7 7.6 6263 1 CALH_MONSE PP3061	7	714		3033	-	POLG_HCVJ8	h genome
9 87 9.1 209 I PAAD_ESEAE 8 8 8.6 452 I AAMP_HUMAN 8 8 8.6 452 I AAMP_HUMAN 8 8 8.6 485 I YI36_TREPA 9 8 8 4 85 I YI36_TREPA 9 8 9 8 4 437 I DEGI_NANTH 7 7 7 9 253 I CAC3_BOVIN 7 4.5 7.8 415 I ZP3_RABIT 7 4.5 7.8 911 I TB11_NEIMB 8 73 5 7.7 730 I HELS_METHAN 7 7 7 7 30 I HELS_METHAN 7 7 7 7 1 1 1 1 1 MANNOSE 7 7 7 6 627 I GRAM_HUMAN 7 7 7 6 627 I GRAM_HUMAN 7 7 7 6 1527 I CALH_MOUSE 7 7 7 6 1527 I CALH_MOUSE 7 7 7 6 1527 I GRAM_HUMAN 7 8 248 I GRAM_HUMAN 7 8 248 I GRAM_HUMAN 7 8 248 I GRAM_MOUSE 7 9 7 6 2 6 6 3 I CENE_HUMAN 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 5 3 2 19 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 5 1 2 10 I GRAM_HOMON 7 1 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	æ	712		3033	~ 1	POLG_HCVJ6	h genome
0 84 9.8 3.21 1 HHOA ARNTH 09scl7 1 82 8.6 485 1 AJAP HUMAN 013685 3 80.5 8.4 437 1 DEGLARTH 022609 4 7.5 7.9 415 1 PAL 1 PAL 7 7.8 415 1 PAL 1 PAL PAL 7 7.8 7.6 1 HYPF_AZOVI P40896 7 7.8 7.6 1 HYPF_AZOVI P40896 7 7.6 1 HYPF_AZOVI P40896 7 7.7 263 1 GRAK_MOUSE 038705 9 7.3 7.7 263 1 GRAK_MOUSE 038705 9 7.3 7.7 7.6 627 1 SADI_MOUSE P51124 1 7.2 7.6 627 1 SADI_MOUSE P51024 1 7.5 248 1 GRAM_MOUSE P305124 1 7.5 248 1 GRAM_MOUSE P305224	o.	87	•	209		PAAD_PSEAE	pscudomon
82 8.6 452 1 AAMP HUANN 013685 82 8.6 452 1 AAMP HUANN 013685 83 8.6 8.6 485 1 1 216_TREPA 022685 84 437 1 DEGL_ARATH 02585 85 74.5 7.8 415 1 ZP3_RABIT PR6883 85 74.5 7.8 75 1 CAC3_BOVIN PR6883 85 74.5 7.8 75 1 CAC3_BOVIN PR6883 85 74.5 7.8 75 1 TP11_NETMB 009056 85 73.5 7.7 263 1 GRAK_HUANN PR6877 85 73.5 7.7 263 1 GRAK_HUANN PR6877 85 75 7 1 CAL1_MONSE 09877 85 75 7 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	10	84	•	321	-	HHOA_ARATH	Q9sel7 arabicopsis
2 8.6 485 1 X136_TREPA 083172 3 8.6 485 1 X136_TREPA 082803 5 7.9 253 1 CAC3_ARATH P02809 5 7.8 415 1 ZP3_RABIT P46833 6 7.4 7.8 911 1 PRIL_BINB P46893 7 7.8 7.6 1 PRIL_BINB 009056 8 7.7 263 1 PRIL_BINB 009056 9 7.3 7.7 730 1 HELS_MCHAR 089777 1 7.2 5.6 257 1 GRAM_HURAN P55124 1 7.2 5.6 257 1 GRAM_HURAN P65124 2 7.6 1527 1 CALH_MONSE P30611 3 7.6 253 1 CENE_HURAN P03124 4 7.6 253 1 CENE_HURAN P03132 5 7.6 253 1 CENE_HURAN P03133 7 7.6 253 1 CEN	11	82		452	r-1	AAMP_HUMAN	
3 80.5 8.4 437 DEGLARATH 022609 4 75 7.9 253 1 CAC3_BOVIN P05805 7 7.8 415 1 PABL P46836 7 7.8 776 1 HYPF_AZOVI P40596 7 7.8 776 1 HYPF_AZOVI P40596 9 73.5 7.7 263 1 GRAK_MOUSE 099056 9 73.5 7.7 263 1 GRAK_MOUSE 035205 1 7.2 7.6 627 1 SADI_MOUSE P5124 2 7.6 627 1 SADI_MOUSE P5124 3 7.6 627 1 SADI_MOUSE P5124 4 7.2 7.6 627 1 SADI_MOUSE P5124 5 7.6 491 1 ERYI_SACER Q02224 7 7.5 248 1 GRAD_MOUSE P10373 7 7.5 248 1 GRAD_MOUSE P10372 7 7.5	12		•	485	1	Y136_TREPA	
4 75 7.9 253 1 CAC3_BOVIN P05805 5 74.5 7.8 415 1 ZP3_RABIT P46833 6 74.5 7.8 911 1 P11_NETMB Q09056 8 7.6 7.6 1 RNE_MOSE Q08076 9 7.3 7.7 263 1 RAM_HUMAN P51124 1 7.2 7.6 627 1 GAM_HUMAN P51124 1 7.2 7.6 627 1 GAM_HUMAN P51124 3 7.7 7.6 1527 1 CAIH_MOSE P3961 4 7.2 7.6 663 1 CENE_HUMAN Q02224 4 7.2 7.6 1848 1 GRAD_MOUSE P3961 7 7.6 18.48 1 GRAD_MOUSE P11033 7 7.5 24.8 1 GRAD_MOUSE P11033 7 7.5 24.8 1 GRAD_MOUSE P11033 7 7.5 24.8 1 GRAD_MOUSE P10	13	O	٠.	437	7	DEG1_ARATH	-
5 74.5 7.8 415 1.273.Rabit P48833 6 74.5 7.8 911 1.811.Methe P40596 7 7.6 7.8 911 1.811.Methe 099056 8 7.7 7.8 1.6 1.871.Methe 099056 9 7.3 7.7 730 1.8LS.Methe 098775 1 7.5 5.6 257 1.8AM.HUMAN 0.8P777 2 7.6 6.27 1.8AM.HUMAN 0.60710 3 7.5 7.6 2.643 1.CENE.HUMAN 0.02244 4 7.2 7.6 2.631 1.CENE.HUMAN 0.02224 5 7.5 2.48 1.RRYI.SACER 0.02224 6 7.1 7.5 3.23 1.VRT.SHRVH P104333 7 7.5 3.23 1.VRT.SHRVH P104715 7 7.5 3.29 1.REALS.CO P03224 7 7.5 3.29 1.REALS.CO	14		٠.	253	~	CAC3_BOVIN	
6 74.5 7.8 776 1 HYPF ALOVT P40996 7 74.5 7.8 911 1 TBIL_NEIMB Q09056 7 7.5 7.8 326 1 PARE_RHIDO Q98056 9 73.5 7.7 263 1 GRAK_MOUSE QS5205 0 73 7.7 263 1 GRAM_HOMSE QS5205 1 72.5 7.6 627 1 SADL_MOUSE P59124 2 72 7.6 627 1 SADL_MOUSE P59124 4 72 7.6 157 1 CALH_MOUSE P59124 5 7.5 263 1 GRAD_MOUSE P59061 7 7.5 7.6 3491 1 ERYL_SACE Q02224 7 7.5 248 1 GRAD_MOUSE P11033 7 7.5 219 1 SPR1_IPOBA P11033 7 7.5 336 1 ULIG_EBV P13MRVH P11033 7 7.5 529 1 PGLZ_RALSO P20041 7 7.5 336 1 ULIG_EBV P13MRVH P03221 7 7.5 529 1 GRAD_MOUSE P11033 7 7.5 7.4 264 1 CTRL_HUMAN P40313 7 7.5 7.4 264 1 CTRL_HUMAN P40313	15	⋖7		415	-	ZP3_RABIT	ory
7 74.5 7.8 911 I TBILL NETHB 009056 8 73 7.8 911 I TBILL NETHB 09956 9 73.5 7.7 263 I GRAK_MOUSE 038675 1 7.2 7.6 1257 I GRAM_HUMAN P51124 1 7.2 7.6 627 I SADL_MOUSE P30671 3 7.2 7.6 623 I CENE_HUMAN P51124 4 7.2 7.6 263 I CENE_HUMAN Q02224 4 7.2 7.6 3491 I ERYL_SACER Q01131 6 7.1.5 7.5 248 I GRAD_MOUSE P11033 7 7.5 248 I GRAD_MOUSE P11033 7 7.5 33 I VENT_SMRYH P11033 7 7.5 336 I ULI6_EBY P11033 7 7.5 336 I ULI6_EBY P03221 7 7.5 336 I ULI6_EBY P0	16	•		176	٦	HYPF_AZOVI	
8 74 7.8 3.26 1 PANE_RHILD 099705 9 7.3 7.7 263 1 GRAK_MOUSE 035205 1 7.2.5 7.6 25.7 1 GRAM_HUNAN 08P777 2 7.6 6.27 1 GRAM_HUNAN 060710 3 7.2 7.6 15.7 1 GALH_MOUSE P39061 4 7.2 7.6 26.3 1 CENE_HUNAN 002224 5 7.5 2.48 1 GRAD_MOUSE P39061 7 7.5 2.48 1 GRAD_MOUSE P10333 7 7.5 2.48 1 GRAD_MOUSE P10333 7 7.5 3.23 1 VPRT_SALPR P10333 7 7.5 3.23 1 VPRT_SALP P10333 7 7.5 3.29 1 ULL6_EBV P10332 7 7.5 3.29 1 ULL6_EBV P10340 7 7.5 3.29 1 CITAL_RUNAN P403323 7 7.4	17	4		911	-	TB11_NEIMB	Q09056 neisseria m
9 73.5 7.7 263 I GRAM_MOUSE 033205 0 73 7.7 263 I GRAM_MOUSE 033205 1 72.5 7.6 627 I SADI_MOUSE P39061 3 72 7.6 627 I SADI_MOUSE P39061 4 72 7.6 5263 I CENE_HUMAN Q02324 5 7.5 7.6 3491 I ERYI_SACER Q03331 6 71.5 7.5 248 I GRAD_MOUSE P11033 7 7.5 7.5 219 I SPRI_IPOBA P11033 7 7.5 336 I ULIG_EBV P21047 9 71 7.5 336 I ULIG_EBV P21047 7 7.5 139 I SPRI_IPOBA P14715 9 71 7.5 139 I TAAI_RAT P18614 7 7.5 7.4 264 I CTRL_HUMAN P14013 7 7.5 7.4 264 I CTRL_HUMAN P14013	18	7		326	-	PANE_RHILO	
0 73 7.7 73.0 I HELS_METHA QBP777 1 72.5 7.6 627 I GRAM_HUMAN P51124 3 72 7.6 627 I SADI_MOUSE Q60710 4 72 7.6 1527 I CALH_MOUSE P39061 5 7.5 248 I GRAD_MOUSE P39061 6 71.5 7.5 248 I GRAD_MOUSE P11033 7 7.5 248 I GRAD_MOUSE P11033 7 7.5 248 I GRAD_MOUSE P11033 7 7.5 33.1 I VPPT_SMRVH P11033 9 71 7.5 33.6 I ULIG_EBV P03221 7 7.5 33.6 I ULIG_EBV P03221 7 7.5 1180 I TAAL_ARA P03221 7 7.4 264 I CTRL_HUMAN P18614 7 7.4 659 I VST2_HEVME Q03500	19	m.		263	-	GRAK_MOUSE	O35205 mus musculu
1 72.5 7.6 257 1 GRAM HUNAN P51124 2 72 7.6 157 1 GALL MOUSE P60710 4 72 7.6 157 1 GALL MOUSE P39061 4 72 7.6 157 1 GALL MOUSE P39061 5 72 7.6 2663 1 GENE HUNAN P60724 5 72 7.6 248 1 GRAD MOUSE P10333 7 7.5 7.5 248 1 GRAD MOUSE P11033 7 7.5 7.5 323 1 VPRT_SHEVH P1033 P1103 7 7.5 7.5 329 1 SPRI_IPOBA P03221 0 7.1 7.5 7.5 1 PGLZ_RALSO P103221 0 7.1 7.5 152 1 PGLZ_RALSO P103221 0 7.1 7.5 1160 1 ITAL_RAT P10614 0 7.1 7.5 160 1 GTRL_HUNAN P40313 0 70.5 7.4 659 1 VSTZ_HEVME P03500	20	~		730	-	HELS_METMA	Q8pzr7 methanosarc
2 7.5 6.27 1.5ADL_MOUSE 060710 3 7.5 7.6 25.7 1.CANL_MOUSE P39061 5 7.6 24.9 1.ERYL_SACER Q03131 6 7.1.5 7.5 24.8 1.GRAL_MOUSE P11033 7 7.1.5 7.5 24.8 1.GRAL_MOUSE P11033 7 7.5 2.4 1.GRAL_MOUSE P11033 7 7.5 2.2 1.VPT_SMRVH P21407 9 7.1 7.5 3.3 1.VLIC_EBV P03221 1 7.5 3.3 1.VLIC_EBV P03221 1 7.5 3.2 1.PGL_RALSO P20041 1 7.5 1.8 1.TAL_RALSO P20041 2 7.4 2.6 1.CTRL_HUMAN P40313 3 7.0.5 7.4 6.5 1.VST2_HEVME Q03500	21	ď		257	-	GRAM_HUMAN	P51124 homo sapien
3 72 7.6 1527 1 CANL, MOUSE P39661-9 4 72 7.6 2463 1 CENE, HUMAN Q02224 7 7.5 248 1 GRAD, MOUSE Q03131 7 7.5 248 1 GRAD, MOUSE P11033 7 7.5 323 1 VPPT_SMRVH P21407 9 71 7.5 336 1 UL16_EBV P03221 7 7.5 336 1 UL16_EBV P03221 7 7.5 529 1 FGL2_RALSO P20041 7 7.5 1180 1 T7A1_RAT P18614 2 70.5 7.4 264 1 CTRL, HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME Q03500	22	72		627	-	SAD1_MOUSE	_
4 72 7.6 2663 1 CENE_HUMAN Q02224 5 71.5 7.6 248 1 GRAD_MOUSE Q03131 7 71.5 7.5 248 1 GRAD_MOUSE P01053 7 71.5 7.5 323 1 VPRT_SHRVH P1033 8 71 7.5 329 1 SPRI_IPPOBA P14715 7 7 7.5 329 1 SPRI_IPPOBA P03221 7 7 7.5 529 1 SGLIZ_RAGSO P20041 7 7 7.5 129 1 FGLIZ_RAGSO P20041 7 7 7.5 7.4 264 1 CTRL_HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME Q03500	23	72	•	1527	. ب	CAIH_MOUSE	_
5 72 7.6 3491 1 ERYL_SACER Q03331 6 71.5 7.5 248 1 GAD_MOUSE P11033 7 71.5 7.5 219 1 VPRT_SMRYH P14715 9 71 7.5 219 1 SPR1_IPOBA P14715 9 71 7.5 336 1 UL16_EBV P03221 1 7.5 180 1 ITAL_RALSO P20041 1 7.5 1180 1 ITAL_HUMAN P40313 2 70.5 7.4 264 1 CTRL_HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME Q03500	24	72	•	2663	-1	CENE_HUMAN	Q02224 homo sapien
6 71.5 7.5 248 1 GRAD MOUSE P11033 7 7.5 7.5 248 1 GRAD MOUSE P11033 7 7.5 7.5 219 1 SPR1_IPOBA P1715 9 71 7.5 336 1 UL16_EBV 7 7 7.5 336 1 UL16_EBV 7 7 7.5 180 1 ITAL_RALSO 2 70.5 7.4 264 1 CTRL_HUMAN 9 70.5 7.4 659 1 VST2_HEVME 9 70.5 7.5 7.5 7.5 7.5 7.5 7.5 7.5 7.5 7.5 7	25	~	٠	3491	-	ERY1_SACER	
7 71.5 7.5 323 1 VPPT_SMRYH P21407 8 71 7.5 219 1 SPR1_IPPBBA P14715 71 7.5 336 1 ULL6_ENV P03221 71 7.5 529 1 FGLIZ_RALSO P20041 71 7.5 1180 1 ITAL_RAT P18614 70.5 7.4 659 1 VST2_HEVME Q03500	56	ï	•	248	_	GRAD_MOUSE	
8 71 7.5 219 1 SFR1_IPOBA P14715 9 71 7.5 336 1 UL16_EBV P03221 0 71 7.5 529 1 FGL2_RALSO P20041 1 71 7.5 1180 1 ITAL_RAT P1864 2 70.5 7.4 264 1 CTRL_HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME 003500	27		٠	323	-	VPRT_SMRVH	P21407 squirrel mo
9 71 7.5 336 1 0116_EBV P03221 0 71 7.5 529 1 PGL2_RALSO P200041 1 71 7.5 1180 1 17A1_RAT P18614 2 70.5 7.4 264 1 CTRL_HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME Q03500	28	71		219	_	SPR1_IPOBA	'n
0 71 7.5 529 1 FOLD.RALSO P20041 1 71 7.5 1180 1 ITAL.RAT P18614 2 70.5 7.4 264 1 CTRL.HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME Q03500	29	71		336	~	UL16_EBV	
1 71 7.5 1180 1 ITAL_RAT P18614 2 70.5 7.4 264 1 CTRL_HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME 003500	30	71	٠	529	П	PGL2_RALSO	P20041 ralstonia s
2 70.5 7.4 264 1 CTRL_HUMAN P40313 3 70.5 7.4 659 1 VST2_HEVME 003500	31	۲,	٠	1180	-	ITA1_RAT	ratt
3 70.5 7.4 659 1 VST2_HEVME 003500	32		٠	264	-	CTRL_HUMAN	~
	33		7.4	629	7	VST2_HEVME) hepa

P19532 homo sapien P06434 human cytom	P06435 human cytom	P28863 oryctolagus	Q8t139 methanosarc	001996 neisseria q	P32434 schizosacch	Q8ye76 brucella me	P05519 bacillus th	P16480 bacillus th	P98160 homo sapien	P08311 homo sapien
TFE3_HUMAN VIE3_HCMVT	/IE2_HCMVT	IMO3_RABIT	IELS_METAC	rBP1_NEIGO	WG2_SCHPO	ONAK_BRUME	4BA_BACTI	CARA_BACTI	PGBM_HUMAN	CATG_HUMAN
	-	~		1	-	7	-	٦	7	-
743 280	410	478	730	915	355	637	1136	1180	4391	255
4.7	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2	7.2
70.5 7.4 7.3										68.5 7.2

ALIGNMENTS

```
CELLUIAR ANINOPEPTIDASE.
CAPASID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E, (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI, PAZZ (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI, PAZZ (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 ROTEWITAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NANA DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SW00487; DEXDC; 1.
Substitution of the process of the polymerase; SNA-directed RNA polymerase; Core procein; Coat procein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MW; 65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GECNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GI.CNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                  InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                     Pfam: PF01006; HCV_NS4a; 1.
Pfam: PF01001; HCV_NS4b; 1.
Pfam: PF01506; HCV_NS5a; 1.
Pfam: PF00271; helicase_C: 1.
Pfam: PF00998; Viral_RdRP: 1.
Probom; PD186062; HCV_NS1; 1.
                                                       HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327197
                                                                                                                                                                                                                      PF01543; HCV_capsid;
PF01542; HCV_core; 1.
PF01539; HCV_env; 1.
PF01560; HCV_NS1: 1.
PF01539; HCV_NS2: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1615
1862
2013
3011
             IPR002519;
IPR002531;
IPR002518;
IPR004109;
                                                                                       InterPro; IPR000745;
InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                  IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3D-structure.
INIT_MET
                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                   Pfam;
SO PRANCE PETER PE
```

```
0
                                                MAPITAYAQQTHGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                          Gaps
                         ..
Length 3011;
                         Indels
                     Ċ
Score 943; DB 1;
Pred. No. 3.4e-82:
                        3; Mismatches
99.0%;
                        Conservative
              Similarity
```

Matches 179:

Local

```
1146 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM 1205
GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRR 120
                                                                                                                                                GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                   01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 23, Last sequence update)
16-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP69) (GP70) (N81); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P65) (P70) (RM-4directed RNA polymerase) (EC 2.7.7.48)].
HS5B (P65) (P70) (RM-4directed RNA polymerase) (EC 2.7.7.48)].
Wiruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-FAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol, 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-9818-321; PUDMed-9483270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Heparlitis C virus NS3 RNA helicase domain with a bound
oligonuclectide: the crystal structure provides insights into the mode
of unwinding:";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -1- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-1- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.

-!-FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.

-!-FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.

-!-FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSEWIAL ROLE IN THE VIRUS REPLICATION.

-!-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thi in Pl and Ser or Ala in Pl'.

-!-CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (RNA)(N)
SUBURIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1
AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genomic structure of the human prototype strain H of hepatitis firus: comparison with American and Japanese isolates.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                 PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A. MEDLINE-92052256; PubMed-1658800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Structure 6:89-100(1998).
                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11108;
                                                                                                                                                                                                                                 181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVH
                                                            61
                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          virus:
                                                                                                                                                                                                                                                                                                                                      RESULT 2
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                    q
                                                                                                                                                                                                                                 ò
                                                          ò
                                                                                                                                                8
```

 $\boldsymbol{\omega}$

```
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1597
1598
1611
1623
1627
1657
1652
1652
AA; 327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           927; DB 1;
No. 1.2e-80;
        (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97.3%;
      1417
1420
1420
1436
11453
11463
11460
11480
11690
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                1312
1313
1323
1343
1352
1350
1352
1358
1373
1376
1376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1533
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
11555
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1382
1389
1397
                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CAR
      CARBOHYD
CARBOHYD
      SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, EI AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ..) (POTENTIAL).
..) (POTENTIAL).
..) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROTEIN P7.
NONSTRUCTURAL, PROTEIN NS2.
PROTEASE/HELLCASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENT I AL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   192
384
384
1052
1052
1052
1052
1053
1063
11065
11065
11065
11065
11065
11065
11065
11065
11065
11065
11065
11065
11065
11065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CCHAIN
CC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                              ö
                                                                                                 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                   09
                                                                                                                                                                                                                      01-APR-1993 (Rcl. 25, Created)
01-APR-1993 (Rcl. 25, Iast sequence update)
01-APR-1993 (Rcl. 25, Iast sequence update)
15-SEP-2003 (Rcl. 42, Iast annotation update)
16-Genome polyprotein (Contains: Capsid protein C (Core protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(Rc 3 4 . 22. -); Proteassc/helicase NS3 (P70) (Hepacivirin)
(RC 3 4 . 21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P66); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NITUSES; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                         GTRT1ASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                  1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A SUBUNIT: THE VIRILOPE CONSISTS OF TWO PROTEINS:
LIPOPROTEIN RAVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
Indels
3;
4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M84754; -; NOT_ANNOTATED_CDS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; S29.001; ...
MEROPS; U39.001; ...
InterPro; IPR001410; DEAD.
Matches 175; Conservative
                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1N64; 25-FEB-03
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=31645;
                                                                                                                                                         RS 1207
                                                                                                                                       RS 182
                                                                                                                                                                                                          POLG_HCVTW
                                                         61
                                                                                                121
                                                                                                                                       181
                                                                                                                                                                                                         a
                                                                                               ò
                                                                                                                                                     셤
                                    g
                                                         ò
                                                                           g
```

```
CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MAJRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1,F22 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                               Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                  Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RELAY SYSTEM (BY
RELAY SYSTEM (BY
RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RELAY SYSTEM
RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 892; DB 1: L
Pred. No. 2.7e-77;
2; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
N-LINKED (GLCNAC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                   ; IPRO02166; HCV_RdRP.
); IPRO07095; RNA_DOL_DS_PS.
); IPRO07095; RNA_DOL_DS_PS.
); IPRO07094; RNA_DOL_PSVIr.
F01543; HCV_core; 1.
PF01542; HCV_core; 1.
PF01560; HCV_NS1; 1.
PF01560; HCV_NS1; 1.
PF01500; HCV_NS2; 1.
PF01006; HCV_NS3; 1.
PF01001; HCV_NS49; 1.
PF01001; HCV_NS49; 1.
PF01005; HCV_NS49; 1.
PF01006; HCV_NS49; 1.
PF01006; HCV_NS49; 1.
PF01006; HCV_NS49; 1.
PF01066; HCV_NS49; 1.
PF01066; HCV_NS49; 1.
PF01067; HCV_NS49; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12;
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS3.
                                                                                              HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2240
2529
2788
                 PR002519;
                                                                                 IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2041
2077
2240
2529
2788
3010
                                                                                                                                                    Pfam; PF01543;
                                                                                                                                                                                                                                    Pfam; PF01006;
                                                                                                                                                                                                                                                Pfam; PF01001;
Pfam; PF01506;
                                                                                                                                                                                                                                                                                          Pfam; PF03998;
                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 164;
                           InterPro;
InterPro;
                                                                                                                           InterPro;
                                                        interPro;
                                                                                                InterPro;
                                                                                                             interPro;
                                                                    InterPro;
                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                         InterPro;
                 nterPro;
                                                                                                                                                                                                                                                                                                        ProDom; F
SMART; SN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
```

ö

Ŋ

```
1026 LAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVÇVVSTATQSFLATCINGVCWTVYHGA 1685
                                                                                                                121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                         GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
1 MAPITAYAOQTRGLLGCTITSLTGRDKNOVEGEVQIVSTAAQTFLATGINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                                                                     01-502 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation update)
66nome polyprotein [Contains: Capsid protein C (Core protein E2
CP68) (GP70) (NS1): Protein P7: Nonstructural protein NS2 (P21)
6EC 3.4.22.-): Protease/helicase NS3 (P70) (Hepacivirin)
6EC 3.4.21.98): Nonstructural protein NS4A (P4): Nonstructural protein
NS4B (P27): Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
Nps5B (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.;
*Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X. FAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088: PubMed-8861916:
LOVE R.A., Parge H.E., Wickcrsham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
Mooraw E.W., Adachi T., Hostomska Z.;
trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998).
-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the vir precursor polyprotein, commonly with Asp or Glu in the P6 possition, Cys or Thr in Pl and Ser or Ala in Pl'.
-!- CATALYTIC ACTIVITY: N uncleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-98227846; PubMcd-9568891; V., Cole J.E., Sardana M., Stan Y., Li Y., Munshi S., Sardana V., Cole J.E., Sardana M., Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.; Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.; Complex of NS3 protease and NS4A peptide of BK strain hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-96235224; Pubbmed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Bur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        from human carriers."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11105;
                                                                                                                                                                                                     1206 RS 1207
                                                                                                                                                                          181 RS 182
                                                                                                                                                                                                                                                                             POLG_HCVBK
P26663;
                                                            61
                                                                                                                                                                                                                                                               POLG_HCVBK
ID POLG_H
                                                                                                                                                                                                                                                                                              a
                                                                                   g
                                                                                                                                                                           à
                        g
                                                          ò
```

```
between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit Anstitutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                           SWISS-PROT entry is copyright. It is produced through a collaboration een the Swiss Institute of Biolnformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSIA (POTENTIAL).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                      PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                  -: SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
InterPro; IPR002221; HCV_capsid.
InterPro; IPR002221; HCV_core.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002518; HCV_NS1.
InterPro; IPR0014109; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS5.
InterPro; IPR001409; HCV_NS5.
InterPro; IPR001099; HCV_NS5.
InterPro; IPR001099; HCV_NS5.
InterPro; IPR001099; HCV_NS1.
IPfam; PF01543; HCV_capsid; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_Core; 1.
Pfam; PF01539; HCV_NS1; 1.
Pfam; PF01500; HCV_NS3; 1.
Pfam; PF01600; HCV_NS3; 1.
Pfam; PF01601; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SMO0487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                           EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                729
1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                               PDB: 1JXP; 14-JAN-98.
PDB: 1NS3: 08-APR-98.
PDB: 1C2P; 15-NOV-00.
PDB: 1GX5; 09-APR-02.
PDB: 1GX5; 09-APR-02.
PDB: 1GX6; 10-APR-02.
PDB: 1QVV; 26-JUN-00.
PDB: 1QVV; 26-JUN-00.
                                                                                                                                                                                                                                                                                                                                                                                          1A10; 25-MAR-98
1JXP; 14-JAN-98
                                                                                                                                                                                                                                                                                                                                                                   PIR; A38465; GNWVTC
PDB; 1A1Q; 25-MAR-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; S29.001;
: U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               347
1083
1107
1165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  116
192
384
730
1007
1616
1863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                 This
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                            01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 23, Last sequence update)
01-40G-1992 (Rel. 24). Last annotation update)
Genome polyprotein (Contains notation update)
Genome polyprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NSI); Protein P7: Nonstructural protein NS2 (PP2);
(EC 3.4.22-); Protease/helicase NS3 (PP0) (Hepacivirin)
(EC 3.4.21-98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RM-directed RNA Polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV)
                                                                                                                                                                                                                                                                       (RNA)(N).
-1- SUBURIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDILINE-9108550; Makagawa M., Ohkoshi S.,
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Molecular cloning of the human hepatitis C virus genome from
"Japanese patients with non-A, non-B hepatitis.",
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                              3010 A.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_core.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ. HSSP; P26663; JJXP.
                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; S29.001; -. MEROPS; U39.001; -. InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR000745;
                                                                                                                                                                                  NCBI_TaxID=11116;
1206 RS 1207
                                            POLG_HCVJA
P26662;
                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                         RESULT 5
POLG_HCVJA
                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 19C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
                                                               (POTENTIAL).
                                                                                                                                                                  (POTENTIAL)
                                     (POTENTIAL)
                                              (POTENTIAL)
                                                                                                    (POTENTIAL)
                                                                                                                               (POTENTIAL)
                                                                                                                                                                                    (POTENTIAL)
                                                                                  (POTENTIAL)
                                                                                          (POTENTIAL)
                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93.5%; Score 691; DB 1; Length 30
89.0%; Pred. No. 3.4e-77;
.ive 15; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327189 MW; F8422D5ECCFDF09C CRC64;
                                                              (GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
(GLCNAC
                                              (GLCNAC.
                                                                                                                                                          (GLCNAC.
                                     (GLCNAC
                                                                                                                                                 (GLCNAC
         N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 89.0%
Les 162; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                   1186
1188
1197
1202
                                                                                                                                                                                                     1047
1050
1063
1074
                                                                                                                                                                                                                                        1081
1081
1085
1087
1092
                                                                                                                                                                                                                                                                                                                         1112
1122
1133
1134
1144
1157
1161
1163
1171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1688
                                                                                                                                                                                                                                                                                              1097
1103
1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3010 AA;
                                                                                                                                                                                             1031
1039
10050
11058
11068
11077
1108
11098
11120
11120
11139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                               CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                          CARBOHYD
                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                        CARBOHYD
CARBOHYD
                                     CARBOHYD
                                                                                  CARBOHYD
                                                                                           CARBOHYD
          CARBOHYD
                                                CARBOHYD
                                                                           CARBOHYD
                                                                                                                                                                                                                                                                  TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                        STRAND
STRAND
STRAND
TURN
STRAND
STRAND
HELIX
TURN
TURN
TURN
TURN
TURN
                                                                                                                                                                                                                                                                                             STRAND
STRAND
TURN
                                                                                                                                                                                                                                STRAND
TURN
STRAND
                                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loca
Matches
                                                                                                                                                                                                                                                           HELIX
qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
61 GTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELORE PROTEIN E ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS! (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         93.1%; Score 887; DB 1: Length 3010; 87.9%; Pred. No. 8.2e-77; Live 17; Mismatches 5; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL
                rpro; IPR001650; Helicase_C.
rpro; IPR007095; RNA_pol_DS_PS.
rpro; IPR007094; RNA_pol_PSvir.
; PF01543; HCV_capsid; 1.
; PF01542; HCV_core: 1.
                                                                                                                                                Pfam: PF01560; HCV_NS1: 1.
Pfam: PF01589: HCV_NS2: 1.
Pfam: PF01006; HCV_NS4: 1.
Pfam: PF01006; HCV_NS4s: 1.
Pfam: PF01001; HCV_NS4s: 1.
Pfam: PF01506; HCV_NS5s: 1.
Pfam: PF0251; Hellcase_C: 1.
Pfam: PF00998; Viral_RGRP: 1.
ProDom: PD186062; HCV_NS1: 1.
IPRO02166; HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3010 AA; 327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          160; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2240
2788
                                                                                                                                                                                                                                                                                                                                                     SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                       InterPro; InterPro;
InterPro;
                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
NP_BIND
SITE
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                        Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
SS STATES STATES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
Medecular cloning of hepatitis C virus genome from a single Japanese cariet: sequence variation within the same individual and among infected individuals.":

Virus Res. 23:39-53(1992).

L. FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
PURDPHOBEC, SUGGESTING A POSSIBLE MEMBRARE-RELATED FUNCTION.
NS3 AND NS5 MAY PARA A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PAYA A ROLE IN THE VIRAL RNA REPLICATION.

L. CATANTYLIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

C -1-CATANTYLIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                           01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-Process (Rel. 42, Last annotation update)
Genome polyprotein (Contains notation and protein (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-JT) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIFOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                           PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_capsid.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS3.
HCV_NS4.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D11168; BAA01943.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Interpro; IPR002522; H
Interpro; IPR002521; H
Interpro; IPR002519; H
Interpro; IPR002519; H
Interpro; IPR002518; H
Interpro; IPR000409; H
Interpro; IPR000465; H
Interpro; IPR000445; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; S29.001;
MEROPS; U39.001;
1206 RS 1207
                                                                                                           POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                        POLG_HCVJT
```

œ

```
đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LAPITAYAQQTRGLLGCIVTSLIGRDKNQVEGEVQVVSTATQSFLATCVNGVCWTVFHGA 1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09
                                                                                                                                                                            CELULAR AMINOBEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GTRIIASPKGPVIQMYINVOKDLVGWPAPQGSRSLIPCICGSSDLYLVIRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MAPITAYAQQTRGLLGCIITSLJGRIJKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Gaps
                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                             SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                         REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          INKED (GLCNAC. . .) (POT 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      92.8%; Score 884; DB 1;
88.5%; Pred. No. 1.6e-76;
                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC)
                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Pred. No. 1.6e
15; Mismatches
                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                     Pfam: PF01001; HCV_NS4b; 1.
Pfam: PF01506; HCV_NS5a: 1.
Pfam: PF00271; helicase_C: 1.
Pfam: PF00998; Viral_RdRP; 1.
Pr0Dom: PD186662; HCV_NS1: 1.
SMART; SM00487; DEXDC: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   326573
        HCV_core; 1.

HCV_env; 1.

HCV_NS1; 1.

HCV_NS2; 1.

HCV_NS3; 1.

HCV_NS48; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 161; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                             1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2529
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RS 182
PF01543;
PF01542;
PF01539;
                              PF01560;
PF01538;
                                                   PF02907;
                                                            PF01006;
                                                                                                                                                               3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                         NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sest Local
                                                  Pfam;
Pfam;
                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
CHAIN
                                Pfam;
  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED custation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                   VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in P1 and Ser or Ala in P1'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
Full-length sequence of a hepatitis C virus genome having poor
homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (RNA)(N).
SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
PROTEIN C AND MENA.
                                                                             3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_core.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A40250; GNWVJ8.
HSSP; P27958; IHEI.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IHEI.
                                                                                                                                                                                                                                                                           NCBI_TaxID=11115;
11
RS 1207
                                                                                                                                                                                                                                                                   Hepacivirus
                                                                             POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                         genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nterPro;
           1206
                                                                 POLC_HCV38
                                                   RESULT
```

```
POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                         Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1090 GNKTLAGPKGPVTQMYISAEGDLVGWPSPPGTKSLDPCTCGAVDLYLVTRNADVIPVRRK 1149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 GTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA 60
                                                                                                                                                                                                                  CAPSID PROTEIN C (POTENITAL).
MAJOR ENVELOPE PROTEIN (POTENITAL).
MAJOR ENVELOPE PROTEIN E (POTENITAL).
NONSTRUCTURAL PROTEIN NSI (POTENITAL).
NONSTRUCTURAL PROTEIN NS2 (POTENITAL).
NONSTRUCTURAL PROTEIN NS4R (POTENITAL).
NONSTRUCTURAL PROTEIN NS4R (POTENITAL).
RNA-DIRECTED RNA POLYMERASE (POTENITAL).
POTENITAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Gaps
                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Score 714; DB 1; Length 3033;
; Pred. No. 3.4e-60;
26; Mismatches 27; Indels C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                              ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        330177 MW;
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1: 1.
Pfam; PF01538; HCV_NS2: 1.
Pfam; PF02907; HCV_NS3: 1.
Pfam; PF01006; HCV_NS4x; 1.
Pfam; PF01001; HCV_NS4x; 1.
Pfam; PF01506; HCV_NS4x; 1.
Pfam; PF01506; HCV_NS5x; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   74.9%;
70.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 70.95
Matches 129; Conscrvative
                                                                                                                                     SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |:
|210 RT 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                     SMART;
                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
POL, HCVUGE

POL,
```

```
Query Match
                                                               PA4019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
      å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTRTIASPKGPVIQMYTNVDKDIJGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                   1 MAPITAYAQQTRGLLGCIIISLTGRDKNQVEGEVQIVSTAAQQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSIEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MW; F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 712; DB 1;
Pred. No. 5.3e-60;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29; Mismatches
                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                           DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PP186652; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          74.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                          SMART; SM00487; DEXDC; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           al Similarity
127; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1210 RS 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RS 182
                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                  FRANSMEN
                                                                                                                                                                                                                                                                                                                                              ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
    οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

209 AM.

PRT;

STANDARD;

PAAD_PSEAE ID PAAD_PSEAE AC Q9HX08;

RESULT 9

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----YLVTRHADVIPVRRRGDSRGSLLSPR--PIS-----YLKGSSGGPLLCPA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8 AQQTRGLLGCIITSLTGRDKNQVEGEVQ-IVSTAAQTFLATCINGVCWTVYHGAGTRTIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               17 AQYGLRLLDCLV-----ETDVA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67 SPKGP------CICGSSDL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HHOA_ARATH STANDARD; PRT; 321 AA.
09SEL7: 049507;
16-0CT-2001 (Rel. 40, Created)
15-SEP-2003 (Rel. 42, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HhoA, chloroplast precursor (EC 3.4.21.-).
HOA OR AT4G18370 OR F28312.39
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                                                                                                                                                                                                                                                                    STRAIN-ATCC 15692 / PAOL;
STRAIN-ATCC 15692 / PAOL;
MEDLINE-Z043737 PubMed-10984043;
MEDLINE-Z043737 PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M. Garber R.L., Goltry L., Tolentino E., Westbrock Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.N., Kas A., Larbig K., Lim R.M., Brody L.L., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of Pseudomonas aeruginosa PAOL, an
                                                                                                                                                                                                               Bacteria; Proteobactería; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PIR; H83144; H83144.
InterPro: IPR003382; Flavoprotein.
Pfam; PF02441; Flavoprotein; 1.
Prothetical protein; Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 WW; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               opportunistic pathogen ";
nature 406:959-964(2000).
-!- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE
PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9.1%; Score 87; DB 1; Length 209; 26.2%; Pred. No. 0.28; iive 18; Mismatches 61; Indels
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               149 GHAVGIFRAAVCTRGVAKAVDFIPVESLET 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --APGFYHQ----PQSVEDLVDFVVARILNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                          Pseudomonadaceae; Pseudomonas.
NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 26.2
nes 55; Conservative
                                                                                                                                                                               Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SSSSEEP SSSS
```

```
Pred
                                                                                                                                    EMBL; AF114386; AAF24060.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   48; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           321 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAMP_HUMAN
013685;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                   TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAMP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                        RA MEDLINE-200834818; PubRed-101/198; RA MEDLINE-200834818; PubRed-101/198; RA MEDLINE-200834818; PubRed-101/198; RA MEDLINE-200834818; PubRed-101/198; RA PODIT. Duoscrenoft A. Stlokema W. Estian K.-D. Terryn N., Hartis B., Amsorge W., Brandt P., Grivell L., Rieger M., Meltler M., de Simone V., Obermaier B. Mache R., Mueller M., Reicherl B., Portetello D., Perez-Alonso M., Boutry M., Bancroft I., Langham S.-A., McCollagh B., Bilham L., Robben J., Andendebussche F., Racken M., Weltjens I. Silham L., Robben J., Andendebussche F., Braken M., Weltjens I., Voet M., Bastlagns I., Andret R., Deffore E., Racken M., Weltjens I., Voet M., Bastlagns I., Andret R., Deffore E., Racken M., Weltjens I., Voet M., Bastlagns I., Andret R., Deffore E., Racken M., Weltjens I., Voet M., Bastlagns I., Koetter P., Braken M., Weltjens I., Voet M., Bastlagns I., Koetter P., Brangel S., Feldpasch II., Hauf J., Koetter P., Brangel S., Feldpasch II., Mullibert M., Deffore E., Racker A., Buysshart C., Gielen J., Villarroel R., Deffore B., Racker A., Buysshart C., Gielen J., Villarroel R., Deffore B., Racker A., Buysshart C., Gielen J., Villarroel R., Deffore B., Racker A., Buysshart C., Gielen J., Willarroel R., Defore B., Racker A., Buysshart C., Gielen J., Willarroel R., Border C., Fuchs M., Farmann B., Granderath K., Dennert T.-H., Doget C., Fuchs M., Farmann B., Granderath K., Dennert T.-H., Doget C., Fuchs M., Valader D., Liquori R., Delaner S., Fancs P., Blenke C., Schnabl S., Hiller R., Schmidt W., Lecharry A., Gasacuberra E., Schnabl S., Hiller R., Schmidt W., Lecharry A., Gasacuberra E., Gibbons T., Webber N., Vandehobl M., Bargues M., Terol J., J., Schwarz S., Scholler P., Globes M., Adbourg S., Francs P., Bennell L., Bedria M., Miller R., Schutz K., Hewes H.-W., Stocker S., Francs P., Bennell L., Bedria M., Shott K., Hewes H.-W., Stocker S., Stoneking T., Kalloki S., Muller R., Schutz M., Marser D., Lamar B., Antoniou B., Zidanic M., Shott K., Doneking T., Mardis E., March M., Bengdoff A., Jones K., Drone K., Cotton M., Johnson 
                                                                                                        Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease.
a homolog to the bacterial periplasmic protease HboA.";
submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CAUTION: Ref.2 sequences differ from that shown due to erroneous gene model prediction. AT4G18370 and AT4G18375 were originally fused into a single gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (JUL-2001) to the SWISS-PROT data bank.
-1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY $2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320. Schubert M., Petcrson U., Funk C., Haas B., Schroeder W.P., Kieselbach T.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        chloroplast lumen from Arabidopsis thaliana.";
II; Brassicales; Brassicaceae; Arabidopsis
                                                                                                                                                                                                                                                                    STRAIN-cv. Columbia;
MEDLINE-20083488; PubMed-10617198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nature 402:769-777(1999).
                                                              [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                      [2]
SEQUENCE FROM N.A.
                      NCBI_TaxID=3702;
```

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 VIPVRRRGDSRG------SLLSPRPISYLK------GSSGGPLLCPA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                218 AI-----GNPYGYENTLTIGVVSGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSY 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  117 LTDEENGKIEGTG-----SGFVWDKLGHIVTNYHVIAKLATDQFGLQRCK 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            59 ----GAGTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam: PF00089; trypsin: 1.
Pfam: PF00089; trypsin: 1.
Pytorisse; Serine profease; Chloroplast: Thylakoid: Transit peptide.
TRANSI 2. CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia: Eutheria: Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 84; DB 1; Length 321;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 LIGRDKNQVEGEVQIVSTAAQTFLATCINGVCW-----TVYH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34691 MW; 68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               149 GHAVGIFRAAVCTR---GVAKAVDF-IPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   273 GHTIGV-NTATFTRKGSGMSSGVNFAIPIDIVVRTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -> G (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROTEASE HHOA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  THYLAKOID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ņ.
                                                                                                                                                                                                                                           EMBL; AL021710; CAA16717.1; ALT_SEQ. EMBL; AL161548; CAB78839.1; ALT_SEQ. MEROPS; SO1.279; -- Interpro; IPR001940; Protease2C. InterPro; IPR001254; Ser_protease_Try. Ffam; PF00089; trybsin; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-GLU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DDT TO DD
```

```
Query Match
Best Local Simi
Matches 44; (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                257
                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                       SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEG1_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCOCCOS SELECT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 WIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSL-----TPCTCGSSDLYLV 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           109 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA------ 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    241 -----VLPDGKRAVVGYEDGTIRIWDLKQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodgon R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.W., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S. Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------GHAVGIFR----AAVCIRGVAKAVDFIPVESL 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEBESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 452;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                           HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           47;
                                                                                                                                                                                                                      MIM; 603488: ...
GO; GO:0008201; F:heparin binding activity; TAS.
InterPro; IPR001680; WD40.
Ffan: PF00400; WD40; 8.
SWART: SM00320; WD40; 8.
PROSITE: PS00678; WD_REPEATS_1; 1.
PROSITE: PS0082; WD_REPEATS_2; 6.
PROSITE: PS50294; WD_REPEATS_2; 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative 13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  485 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hypothetical lipoprotein TP0136 precursor TP0136.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8.6%; Score 82;
25.3%; Pred. No. :
                                                                                                                                                                                                                                                                                                                                                                                                                          POLY-GLU.
WD 1.
WD 2.
WD 3.
WD 4.
WD 5.
WD 6.
WD 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              49015 MW:
                                                                                                                                                           EMBL; M95627; AAA6889.1; -, PIR; I39383; I39383.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                         18
77
138
180
220
261
306
363
                                                                                                                                                                                               Genew; HGNC:18; AAMP.
                                                                                                                                                                                                                                                                                                                                                                                     Repeat: WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            452 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity
Matches 42; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Treponema pallidum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Y136_TREPA
083172;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                             NIAMOO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL cutstains. the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78 NVDKDLVG-----WPAPQGSRSLIPCTC-----GSSDLYLVTRHADVIP-----VR 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23 IGRDKNQVEGEVQIVSTAAQTFLATCI -- NGVCWTVYHGAG---TRTIASPKGPVIQMYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141 T----LVGGTSRPFWLVPGGTGNNGNCGCGGGGGSSSSSSCIHIWLVPGGTGNNGNCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
spirochete.";
Science 281:375-388(1998).
-!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDIINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam 2.;
Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
HYPOTHETICAL LIPOPROTEIN TP0136.
N-ACYL DIGLYCERIDE (POTENTIAL).
                                                                                                 (Potential).
-!- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8.6%; Score 82; DB i; Length 485; 24.2%; Pred. No. 2.2; tive 13; Mismatches 65; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEGLARATH STANDARD; PRT; 437 AA.
022609; O9LR85;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
BEGP1 OR DEGP OR AT3627925 OR X16N12.18.
Arabidopsis thallana (Wouse-ear cress)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; AE001199; AAC65137.1; ALT_INIT.
TIGR; TP0136; --
Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLY/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLY/SER-RICH.
GLY/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            48984 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. STRAIN-CV. Columbia;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              444
485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VG 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GG 258
```

```
210 PK--NKIRPIPVGVSADILVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S1
                                                                                                                                                                                                                      (Rel. 09, Created)
(Rel. 37, Last sequence update)
(Rel. 42, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- SUBCELLULAR LOCATION: EXTRACEllular.
-!- TISSUE SPECIFICITY: Pancreas.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY (PDB; 1FON; 14-OCT-96.
PDB: 1PY; 27-JAN-97.
MEROPS; S01.983; -.
                                                                                                                                                                                   253 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [2]
SEQUENCE OF 14-253, AND DISULFIDE BONDS.
MEDLINE-86220198; PubMed-3319215;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro: IPR001314; Chymotrypsin.
InterPro: IPR001254; Ser_protease_Try.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF00089; trypsin; 1.
PRINTS: PR00722; CHYMOTRYPSIN.
SMART; SM00020; Tryp_SPc: 1.
PROSITE; PS50240; TRYPSIN_DOW; 1.
PROSITE; PS00134; TRYPSIN_HIS; 1.
PROSITE; PS00135; TRYPSIN_HIS; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eur. J. Biochem. 157:91-99(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            procarboxypeptidase A compley
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Serine protease homolog;
PROPEP 1 11
                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       253
57
103
206
187
227
                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=9913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pulgserver A.;
                                                                                                                                                                                                                    01-NOV-1988 (15-DEC-1998 (15-SEP-2003 (
                                                                                                                                                                                 CAC3_BOVIN
P05805;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRRGDSRGSLLSPRPI 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --- TNVDKDLVGW-----PA 88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSGLSNST -> SSTLFLHSPPSSHL (IN REF. 2).
V -> I (IN REF. 2).
P -> S (IN REF. 2).
G -> R (IN REF. 2).
L -> FF (IN REF. 2).
L -> YF (IN REF. 2).
L -> V (IN REF. 2).
L -> V (IN REF. 2).
L -> V (IN REF. 2).
C -> D (IN REF. 2).
C -> D (IN REF. 2).
C -> C (IN REF. 2).
                                                                                                                                                                                                                                                                           ENZYME REGULATION: INHIBITED BY PHENYLMETHYL, SULFONYL, FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                       STRAIN=cv. Columbia;
Kieselbach T., Bystedt M., Schroeder W.P.;
Submitted (Jul.-2000) to the SWISS-PROT data bank.
-i- FUNCTION: SERIE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
CAN DEGRADE BETA-CASEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ransit peptide; Chloroplast; Thylakoid.
CHLOROPLAST (POTENTIAL).
                    Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
Structural analysis of Arabidopsis thaliana chromosome 3. 11.
Sequence features of the 4.251,695 bp regions covered by 90 Pl, TAC and BAC clones.";
                                                                                                                                                                                                                                                                                                              -1 - SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8.4%; Score 80.5; DB 1; Length 437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       _
__
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1497BIAB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                 -1- INDUCTION: By heat shock.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-1- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROTEASE DO-LIKE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SERINE PROTEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001940; Protease2C,
InterPro; IPR001254; Ser_protease_Try.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AP000371; BAB02539.1; -
EMBL; AP001302; BAB02539.1; JOINED.
MEROPS; S01.279; -
MEDLINE-20363099; PubMed-10907853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AF028842; AAC39436.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56 VYHGAGTRTIASPKGPVIQMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRINTS; PR00834; PROTEASES2C.
SMART; SM00228; PDZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE; PS50106; PDZ; 1.
Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00595; PDZ; 1.
Pfam; PF00089; trypsin; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001478; PDZ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative
                                                                                                  DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           103
437
321
421
171
201
280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36
54
54
64
64
63
355
355
                                                                                                                    [3]
SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                              O-PHENANTHROLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          416
437 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity
Matches 44; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                104
152
324
171
201
                                                                                                                                                                                                                                                                                                                                  MEMBRANE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
"Autolysis of proproteinase E in bovine procarboxypeptidase A ternary complex gives rise to subunit III."; FEBS Lett. 277:37-41(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Proproteinase E precursor (Procarboxypeptidase A complex component III) (Procarboxypeptidase A-S6 subunit III) (PROCPA-S6 III).
                                                                                                                                                                                                                                                                                                                                                            Bos raurus (Bovine).

Bukaryota Metazoa, Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos
134 SYL-----KGSSGGPLI.CPAGHAVGIFRAAVCTRCVAKAVDF-IPVESL 176
                                                 266 QDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGASSGVGFSIPVDTV 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Venot N., Sciaky M., Puigserver A., Desnuelle P., Laurent G., "Amino acid sequence and disulfide bridges of subunit III, a defective endopeptidase present in the bovine pancreatic 6 S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [1]
SEQUENCE OF 1-25.
MEDLINE-91099520; PubMed-2269366;
Pascual R., Vendrell J., Aviles F.X., Bonicel J., Wicker C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pancreas; Digestion; 3D-structure.
ACTIVATION PEPTIDE.
PROPROTEINASE E.
```

```
Search completed: August 30, 2003, 19:13:43
Job time : 13.0138 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                44987 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                         18
415
378
399
415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                415 AA;
                                                                                                          SEQUENCE FROM N.A.
                                                                                    NCB1_TaxID-9986;
                                      ZPC.
                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100 -CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAG-----HAV 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                170 HCSQWDWWGITVKKTM--VCAGGDTR-----SGCNGDSGGPLNCPAADGSWQVHGV 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----SPKGPVI-----OMYTNVDKDLVGWPAPQG-SRSLTPCT--- 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17 CIITSLIGR-----DKNQVEGEVQIVS-TAAQIFLATCINGVCWIVYHGAGIRTIA---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1.9%; Score 75; DB 1; Length 253; llarity 22.7%; Pred. No. 4.9; Conservative 23; Mismatches 72; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27337 MW; 24663724D8AE409C CRC54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (Rel. 33, Created)
(Rel. 33, Last sequence update)
(Rel. 33, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              415 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GIFRAAVCTRGVAKAVDFIPVES 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 219 TSFVSAFGCNTIKKPTVFTRVSA 241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                            Query Match
Best Local Similarity
Matches 46; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      253 AA;
                                                                                                                                                                                                        11136
11136
11136
11136
11136
11136
11136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
1136
ZP3_RABIT
P48B33;
01-FEB-1996 (
01-FEB-1996 (
01-FEB-1996 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HELIX
HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           67
                                                                                                                   STRAND
STRAND
STRAND
TURN
STRAND
HELIX
                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
2P3_RABIT
ID 2P3_R
AC P4883
DT 01-FE
DT 01-FE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ς;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55 TVYHGAGTRIIASPKGP-VIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHAD 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families."; DNA Seq. 4:361-393(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 VIPVRRRGDSRGSLLSPRPISYLKGSSGGPL--LCPAGHAVGIFRAAVCTRGVAKA 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels 15; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20NA PELLUCIDA SPERM-BINDING PROTEIN EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                               TISSUE-Ovary;
MEDLINE-95143578; PubMed-7841460;
Harris J.D., Hibler D.W., Fontenot G.K., Hsu K.T., Yurewicz E.C.,
                                                                                                                               Oryctolagus cuniculus (Rabbit).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
2ona pellucida sperm-binding protein 3 precursor (2ona pellucida
qlycoprotein 2P3) (Sperm receptor) (2ona pellucida protein C)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PSOG682; ZP_DOMAIN; 1.
Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
Extracellular matrix; Multigene family.
NON.TER 1 18 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 415;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          77396CF1BAA3F5CB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.8%; Score 74.5; DB 1; Illarity 26.7%; Pred. No. 9.5; Conservative 17; Mismatches 53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL: U05782; AAA74392.1; -.
PIR: S70401; S70401.
Interpro: IPR001507; Endoglin/CD105.
Pfan: PF00100; zona_pellucida; 1.
SMART; SM00241; ZP; 1.
```

Н

000000000000000000000000000000000

hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis

Run ĕ

```
2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-PL.1Y;
HOlland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Gonetic Diversity and response to IFN of the NS3 Protease Gene from
clinical Strains of the Hopatitis C Virus.*;
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases,
EMBL. ARK84560.1;
-1. InterPro; IPR004109; HCV_NS3.
Piam: PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                          no DNA stage; Flaviviridae;
                                                                       091rr2
091rr7
091rr7
091rr1
091rs1
091rs2
091rs0
091rr8
091rr8
091rr8
036609
                   091rr5
091rs4
091rt6
091rt9
091rr4
091rr9
                                                                                                                                                                                                                                                                                                                                                                                                              99.2%; Score 945; DB 12; Length 181; larity 99.4%; Pred. No. 1.9e-88; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, Hepacivirus.
                                                                                                                                                                                                                                                         Created)
                                                    091RR9
091RR0
091RR2
091RR2
091RT2
091RT2
091RT2
091RR1
091RR3
091RR3
                                                                                                                                  09PWU9
091RS6
091RT7
091RS0
091RT8
                                                                                                                                                                                                                                          O91RR8 PRELIMINARY;
O91BC-2001 (TERMELRE) 19, Ci
01-DEC-2001 (TERMELRE) 19, La
01-DEC-2001 (TERMELRE) 29, La
NS3 protease (Fragment).
Hepatitis C virus.
Best Local Similarity
Matches 180; Conserv
                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                  091rr8 hepatitis c 091rr5 hepatitis c 091re5 hepatitis c 091rs hepatitis c 091rs hepatitis c 091rq8 hepatitis c 091rs hepatitis c
                                           August 30, 2003, 19:00:22; Search time 34.7298 Seconds (without alignments) 1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                   1 MAPITAYAQQTRGLLGCIIT.......GVAKAVDFIPVESLETTMRS 182
                                                                                                                                                                                                                                                                                                                                                                        Description
     5.1.6
Compugen Ltd.
                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                   830525 seqs, 258052604 residues
     GenCore version
Copyright (c) :993 - 2003
                                                                                                                                                                                summaries
                                - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                   091848
091875
0918756
0918765
091881
091881
091889
091889
091889
                                                                                                       Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                        sp_unclassified:*
                                                                                                                                                                                                                                                             spplant:*
sp_rodent:*
sp_virus:*
sp_vertebrate:*
                                                                                                                                                                                                                              sp_invertebrate:*
                                                                                                                                                                         Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                     sp_bacteriap:*
                                                                                                                                                                                                                                                sp_organelle:*
sp_phage:*
                                                                                                                                                                                            SPTREMBL_23:*
: sp_archea:*
: sp_bact.eria:*
                                                                                                                                              seq length: 0 seq length: 20000000000
                                                                                                                                                                 Post-processing: Minimum Match 0%
                                                                                                                                                                                                                                                                                                                                                                                     sp_rvirus:*
                                                                                                                                                                                                                                                                                                             sp_archeap:
                                                                                                                                                                                                                                                                                                                                                                      103
                                                                                                                                                                                                                                    sp_mammal:*
                                                                      US-09-965-594-1
                                                                                                                                                                                                                Sp_fungi:*
Sp_human:*
                                                                                                                                                                                                                                                                                                                                                                Query
Match Length
                                                                                                                                                                                                                                           sp_mhc:*
                                                                                                 BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                              တက္ထာထားတယ္က
                                                                                                                                                                                                                                                             10:
                                                                                                                                                                                                                                                                                                                                                                                   Perfect score:
                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                      Score
                                                                                                                                             Minimum DB s
Maximum DB s
                                protein
                                                                                                                    Searched:
                                                                                    Sequence:
                                                                                                                                                                                            Database
                                             on:
                                                                                                                                                                                                                                                                                                                                                                                     65433210987654321
                                                                                                                                                                                                                                                                                                                                                                 Result
```

121 120

9

61

Gaps

ö

```
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9IFES
Q9IFES;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                9
DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 APITAYAQQTRCLI.GCIIISLTGRDKNQVRCEVQIVSTAAGFFI.ATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 IRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-Pt.4;
Holland-Steley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Steley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatilis C Virus.*;
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF869218; AAK54543.1;
InterPro; IPR004109; HCV.NS3.
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 945; DB i2; Length 181;
Pred. No. 1.9e-88;
1; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. Choo Q.-L., Richman K., Han J.; "The nucleotide sequence of the Hepatitis C viral genome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q81756 PRELIMINARY; PRT; 2436 AA.
Q81756;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                           181 AA.
                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tches 180; Conservative
                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S 181
                                                                                                                           182 S 182
                                                                                                                                                                                      5 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
122
                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                        Q91RT5
                                                                                                                                                                                                                                                                                                                                           Q91RT5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 3
(9176
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
(10 26)
                                                                                                                                                                                                                                                                                 RESULT 2
Q91RT5
                                                             셤
                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                              NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
ð
```

```
61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [1] SEQUENCE FROM N.A. SEQUENCE FROM N.A. MEDINE-21262212; PubMed-11369872; Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.; Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.; Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepatuirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99.0%; Score 943; DB 12; Length 2436; 98.4%; Pred. No. 7.1e-87; tive 3; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2436 2436
2436 AA; 264734 MW: D7B9872900BE3125 CRC64;
EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-0cT-2000 (TrEMBLrel. 15, Created)
01-0cT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001490; HCV_NS4D.
InterPro; IPR002868; HCV_NS5a.
InterPro; IPR001265; HCV_RGRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sequence.";
J. Gen. Virol. 82:1291-1297(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pram; PF01560; HCV_NS1; I.Pfam; PF01560; HCV_NS1; I.Pfam; PF01300; HCV_NS3; I.Pfam; PF01006; HCV_NS4a; I.Pfam; PF01006; HCV_NS4b; I.Pfam; PF01506; HCV_NS4b; I.Pfam; PF00271; helicase_C; I.Probom; PF0098; Viral_RGRP; I.Probom; PF0098; Viral_RGRP; I.SMART; SMO0487; DEXDC; I.
                                                                                                                                                                                                                                                    IPR004109; HCV_NS3.
IPR000745; HCV_NS4a.
    the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 98.4
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                    IPR002518;
IPR004109;
(MAY-1990)
                                                                                                                                 InterPro; IPR001410;
                                                                                                                                                                          IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11
RS 757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
(9) FES
10 (9) FES
20 (9
```

ö

```
Best Local Similarity 98.9
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 98.9
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                           SEQUENCE FROM N.A.
                                NCBI_TaxID-111.03;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 S 181
        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                       Protease.
                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  091RS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
Q91RS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | InterPro; | IFR002166; | HCV_ACADO. | |
| InterPro; | IFR002166; | HCV_ACADO. |
| InterPro; | IFR001050; | HCV_ACADO. |
| InterPro; | IFR001050; | HCV_ACAD. |
| InterPro; | IFR007094; | RNA_DOI_DS_PS. |
| InterPro; | IFR007094; | RCV_CAD.; | I. |
| InterPro; | IFR007094; | RCV_CAD.; | I. |
| InterPro; | IFR007094; | RCV_CAD.; | I. |
| InterPro; | IFR007096; | RCV_NS3; | I. |
| InterPro; | IFR007096; | RCV_NS3; | I. |
| InterPro; | ITRO. |
| InterPro; | ITRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NG3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED HY LIPOPROTEIN BUNELOPE. CONSISTS OF TWO PROFINS: PROFIEN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROFIEIN C. AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ċ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 943; DB 12; Length 3011; Pred. No. 9.2e-87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase: Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase: Transferase: Transmembrane.
SEQUENCE 3011 AA; 327124 WW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                         InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR001410; DRAD.
                                                                                                                                                                                                                           InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_capsid.
InterPro: IPR002521; HCV_env.
InterPro: IPR002531; HCV_NS1.
InterPro: IPR004109; HCV_NS2.
InterPro: IPR004109; HCV_NS3.
InterPro: IPR000745; HCV_NS4.
InterPro: IPR001490; HCV_NS4.
InterPro: IPR001490; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3;
                                                                                                                   EMBL; AF271632; AAF81759.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99.08;
98.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091RR3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q91RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091RR3
1D 09
AC 09
DT 01
DT 01
DE NS
  SON KEEN TO DEED THE TO THE TOTAL TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 60
                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                            2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-PL.K;
STRAIN-PL.K;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.';
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AF868232; AAK54557.1;
InterPro: IPR004109; HCV_NS3.
Protease.
STRAIN-FL.4B:
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Hondrad-Strains of the Heperitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AF869240; ARK54565.1:
InterPro: IPR004109; HCV_NS3:
Pfam; PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 181;
                                                                                                                                                                                                                                                                                      Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                      181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 181
181 AA; 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                            . ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98.9%; Score 942; DB 12;
98.9%; Pred. No. 3.8e-88;
Live 2; Mismatches 0;
                                                                                                                                                                                                                                                                                      Score 942; DB 12;
Pred. No. 3.8e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA
                                                                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                      98.84;
98.94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23,
```

```
SEQUENCE FROM N.A.
  [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-Pt. 3T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                       NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                    Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q91RR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
091RR6
    ò
                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCOCCOS SELECT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ပဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61 TRIIASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCTCGSSDLYLVTRHADVIPVRRG 120
                                                                                 122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTWR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVOIVSTAAQTFY,ATCINGVCWTVYHCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCICGSSDLYLVTRHADVIPVRRRG
TRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.J.;
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF856245; AR854570.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98.8%; Score 942; DB 12; Length 181; 98.9%; Pred. No. 3.8e-88; Live 2; Mismatches 0; Indeis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 181
181 AA: 19144 MW; COC91F1E2EEB0B32 CRC64;
                                                                                                                                                                                                                                                                                                                                      01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-DEC-2001 (TIEMBLrel. 19, Created)
01-DEC-2001 (TIEMBLrel. 19, Last sequence update)
01-MAR-2003 (TIEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 AA
                                                                                                                                                                                                                                                                                                    181 AA
                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                 NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
Matches 179; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                             182 S 182
                                                                                                                                                                                             s 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Pt.52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182
                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091RT1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
091RT1
1D 091RR
AC 091RR
DT 01-D
DT 01-D
DT 01-D
OC WITU
OC Hepa
                                                                                                                                                                                                                                                                            õ
                                                                                                                 g
                                                                                                                                                           ò
                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                        a
ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 TRIIASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCICGSSDLYLVTRHADVIPVRRG 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRIIASPKGPVIQMYINVDKDLVGMPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 TRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to Irw of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865221, AAR5447.1:
Interpro; IPR004109; HCV NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L., Genetic Diversity and response to IRN of the NS3 Protease Galinical Strains of the Hepatitis C Virus ". Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL, RES89237. AARS4562.1; --
InterPro: IPRO4109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 98.6%; Score 940; DB 12; Length 181; Best Local Similarity 98.3%; Pred. No. 6e-88; Matches 178; Conservative 3; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                     Length 181;
                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 1
181 181
181 AA; 19101 MW; 614ADABBOF33CCAF CRC64;
                                                                                                                                                                                                     1 1
181 181
181 AA; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                        .;
0
                                                                                                                                                                                                                                                                                                  Query Match 98.8%; Score 942; DB 12; Best Local Similarity 98.9%; Pred. No. 3.8e-88; Matches 179; Conservative 2; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091RR6;
01-DEC-2001 (TrEMBLrel. 19, C)
01-DEC-2001 (TrEMBLrel. 19, L,
01-MAR-2003 (TrEMBLrel. 23, L,
NS3 protease (Fragment)
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
```

S

유 à g à g

```
122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-HC-JI; MEDLINE-9204440; PubMed-1658196; MEDLINE-9204440; PubMed-1658196; MEDLINE-9204440; Machida A., Okadoto H., Mavumi M.; Miyakowa Y., Mavumi M.; Mucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier. comparison with reported isolates for conserved and divergent regions. "." Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                           1 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          62 TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG
                                                                                                                                                                                                                                                                                                                                                                                      2 APITAYAQQIRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGAG
                                            "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (ASR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL: AFS69230: AAK54555.1;
InterPro; IPR004199; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDILINE-3317120; PubMed-1335573;
Okamoto H., Kanai N., Mishiro S.;
"Full-Inength nucleotide sequence of a Japanese hepatitis C virus isolate (HC-11) with high homology to USA isolates.";
Nucleic Acids Res. 20:6410-6410(1992).
                                                                                                                                                                                                                                                                                                                                          ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "The 5'-terminal sequence of the hepatitis C virus genome."; Jpn. J. Exp. Med. 60:167-177(1990).
                         Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
                                                                                                                                                                                                                                                                                                  Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                     1 1
181 181
181 AA; 19132 MW; 0BB90B5F3AB95250 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                             98.5%; Score 939; DB 12; 98.3%; Pred. No. 7.6e-88; iive 3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-HC-J1;
MEDLINE-91013116; Pubmed-2170712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01,
01-NOV-1996 (TrEMBLrel. 01,
01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                Local Similarity 98.3
les 178; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S 181
    STRAIN-Pt. 24;
                                                                                                                                                                                                       NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q03463
Q03463;
                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      003463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q,
                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 DSRGSILLSPRPISYLKGSSGGPLICPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMR 180
                                                                   DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
61 TRIIASPKGPVIQMYINVDQDLVGWPAPQGARSLIPCTCGSSDLYLVTRHADV1PVRRRG 120
                                              DSRGSLLSPRP1SY1,KGSSGGP1,LCPAGHAVGIFRAAVCIRGVAKAVDFIPVESLFTTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 APITAYAQQTRGLLGCIIISLIGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         from
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.; "Genetic Diversity and response to IrN of the NS3 Protease Gene Clinical Strains of the Hepatitis C Virus."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL; AR586249.1: ---
Interpro; IPR004109: HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 12; Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 AA: 19131 MW; 8BD7FC2769DBD635 CRC64;
                                                                                                                                                                                                                                                                                                                                   01-DEC-2001 (TIEMBLIEL. 19, Created)
01-DEC-2001 (TIEMBLIEL. 19, Last sequence update)
01-MAR-2003 (TIEMBLIEL. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98.6%; Score 940; DB 12;
98.9%; Pred. No. 6e-88;
Live 2; Mismatches 0;
                                                                                                                                                                                                                                                                                             181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091RS3;
01-DEC-2001 (TEMBLEEL 19,
01-DEC-2001 (TEMBLEEL 19,
01-MAR-2003 (TEMBLEEL 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 98.9
Matches 179; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                    NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                              S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             s 181
                                                                                                                                    S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Protease.
                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122
                                              122
                                                                                                                                    182
                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RS3
                                                                                                                                                                                                                                        RESULT 10
091R59
10
091R5
10
091R5
10
01-DE
DT
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
091RS3
AC 091RS
AC 091RS
DT 01-DE
DT 01-DE
DT 01-MB
DE NS3 P
OC HEPAC
OC HEPAC
OX NCEI
RN (1)
RN (1)
```

ò g ö 셤 ò g ò

121

us-09-965-594-1.rspt

```
NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                         Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q91RS8
Q91RS8;
                                       091RT4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ö,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA
                                                                                                                            "Characterization and mapping of a B-cell immunogenic domain in hepatitis C virus E2 glycoprotein using a yeast peptide library."; Virology 200:246-255(1994).
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: D10749: BAA01582.1; -- HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SMANT; SMOVENCY, DEALLY, PROBE POSITIVE; 1.
PROSITE; PS50501; RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 98.5%; Score 939; DB 12; Length 3011; Best Local Similarity 97.8%; Pred. No. 2.4e-86; Matches 178; Conservative 3; Mismatches 1; Indels 0
                                                                                                      Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97E9052C0250463B CRC64;
                                       to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
InterPro; IPR001522; HCV_capsid.
InterPro; IPR001522; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR00219; HCV_NS2.
InterPro; IPR001218; HCV_NS2.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR00166; HCV_RGAB.
InterPro; IPR00166; HCV_RGAB.
InterPro; IPR001016; HCV_RGAB.
InterPro; IPR001059; RNA_pol_DS_PS.
InterPro; IPR001099; RNA_pol_DS_PS.
InterPro; IPR001099; RNA_pol_DS_PS.
InterPro; IPR001099; RNA_pol_PSvir.
Pfam; PF01542; HCV_core; J.
Pfam; PF01542; HCV_Core; J.
Pfam; PF01506; HCV_NS2; I.
Pfam; PF01006; HCV_NS3; I.
Pfam; PF01006; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Pfam; PF01001; HCV_NS4b; I.
Probom; PD186062; HCV_NS1; I.
PROBOM; PD186062; HCV_NS1; I.
                                                                                           MEDLINE-94174722; PubMed-7510436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327112 MW;
                             Okamoto H.;
Submitted (DEC-1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3011 AA;
                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RS 182
               STRAIN-HC-J1;
                                                                                                                     Inchauspe G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 TRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AS369225; AAK54550.1;
InterPro: IPR004109; HCV_NS3;
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.

STRAIN-Pt.23;

Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;

Holland-Strains of the Reponse to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus. ;

Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL; AF369219; AAR45444.1;

InterPro; IPR004109; HCV_NS3;

Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Plaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98.4%; Score 938; DB 12; Length 181; 98.3%; Pred. No. 9.6e-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 1
181 181
181 Aa; 19059 MW; 1E53C47AE8B7E5C9 CRC64;
                                                                          01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AA
181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 9.6e
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 178; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Pt. 176;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182 S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAC DOT REPORT OF THE REPORT O
```

```
Job time : 36.7298 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ပ
                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG 121
                                                                                                                                                                                                                                                                                                            DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMR 181
                                                                                                                                                                                                                                                                                 62 TRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           90
                                                                                                                                                                                                       5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 APITAYAQQTRGLLGC11TSLTGRDKNOVEGEVQ:VSTAAQTFLATC1NGVCWTVYHGAG
                                                                                                                                                                               2 APITAYAQOTRGLIGGIITSIJGRDKNOVEGEVOIVSTAAQTFLATCINGVCWTVYHGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-201) to the EMBL/GenBank/DDBJ databases.
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                               ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 181;
                                                                               Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                            Indeis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 181
181 AA: 19116 MW; 96488G7F49EBID43 CRC64;
                          181 AA; 19114 MW; 574AC47AE8AEE5D2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98.4%; Score 938; DB 12; 98.3%; Pred. No. 9.6e-68;
                                                                          Score 938; DB 12;
Pred. No. 9.6e-89;
3; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                            3,
                                                                          98.48;
98.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 178; Conservative
                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [1]
SEQUENCE FROM N.A.
                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      S 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-Pt.11
                                                                                                       Best Local Sim
Matches 178;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                          SEQUENCE
                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q91RT3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
091RT3
1091RT4
AC 091RT
DT 01-DE DT 01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                              Ω
F 08
                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

Search completed: August 30, 2003, 19:18:17

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

using sw model protein search, , OM protein

August 30, 2003, 19:02:22 ; Search time 14.9785 Seconds (Without alignments) 1168.452 Million cell updates/sec Run on:

US-09-965-594-1 953 Title: Perfect score: Sequence:

1 MAPITAYAQQTRGLLGCIII.........GVAKAVDFTPVESLETIMRS 182

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 segs, 96168682 residues Searched:

283308 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_76:* Database :

pirl:*
pir2:*
pir3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ES	Description	genome polyprotein	genome polyprotein				genome polyprotein		genome polyprotein	genome polyprotein		Ď.	ാ	1	probable aromatic	angio-associated m	hypothetical prote	heat shock transcr	cytochrome-c oxida	beta transducin ho	serine proteinase	probable periplasm	hypothetical prote	hypothetical prote	aconitate hydratas	procarboxypeptidas	zona pellucida gly	hypothetical prote	aconitate hydratas	transferrin-hindin
SUMMARIES	CII	GNWVC3	S40770	GNWVCH	CNWVTW	GNWV7C	GNWVCJ	A45573	S18030	JC5620	GNWVJ8	J01303	T08841	T08839	1183144	139383	B71360	T48399	F83976	T18234	н89966	B71284	A84212	T06118	H81775	CPBOA3	S70401	S23441	CB1200	JN0821
	DB	-	-	٦	٦	~	7	Н	٦	Н	r-1	-	7	~	7	~	~	7	ď	7	~	7	7	7	~	-	~	~	7	~
	Ouery Match Length		3011	3011	3010	3010	3010	3010	3010	3014	3033	3033	3005	2970	209	452	495	476	620	981	239	398	904	270	898	240	415	755	868	911
æ	Ouery Match	99.0		۲.	'n.	'n.	93.1	92.8	90.7	÷	74.9	4.	28.1	26.8	9.1	8.6	8.6	9.6	8.5	8.5	8.5	8.1	8.1	0.8	•			7.8		7.8
	Score	943	939	927	892	891	887	884	864	804	714	712	267.5	255.5	87	82	82	81.5	78.5	78.5	78	77.5	77	76.5	76.5	^	4	74.5	Ψ.	74.5
	Result No.		7	e	❖	S	9	7	æ	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

hypothetical prote	hypothetical prote	protein F22G5, 28 1	hypothetical prote	hypothetical prote	C4-dicarboxylate t	phosphate ABC tran	hypothetical prote	transferrin-bindin	transferrin-bindin	latent transformin	serine proteinase	hypothetical prote	htrA-like serine p	chitinase (EC 3.2.	granzyme M (EC 3.4
C72705	A75335	G86208	F90978	S76033	H82044	H84203	T04533	C81832	F81196	A55494	A89967	T49806	H97199	565762	A55634
7	~	~	7	7	a	~	~	7	7	7	7	~	7	~	7
140	377	808	356	447	451	266	846	910	915	1820	239	354	433	535	249
7.8	7.8	7.8	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.7	7.6
74	74	74	73.5	73.5	73.5	73.5	73.5	73.5	73.5	73.5	73	7.3	73	73	72.5
30	31	32	33	34	35	36	37	38	36	0.7	4.1	42	43	44	45

ALIGNMENTS

```
A; Responses: 1977-1033 CHAA.
A; Cross references: DBBJ:D10128
A; Experimental source: isolates E-b16
A; Accession: PQ0404
A; Scatus: preliminary
A; Accession: Source: Isolates E-b17
C; Superfamily: hepatilis C virus genome polyprotein
C; Superfamily: hepatilis C virus genome polyprotein; glycoprotein; hydrolase; nonstru
C; Superfamily: hepatilis C virus genome polyprotein
F; 1-115, Product: angle protein C *status predicted cMED>
F; 116-115, Product: monstructural protein E *status predicted cMED>
F; 390-729/Product: nonstructural protein NS2 *status predicted cMS2>
F; 1007-1615/Product: hepatilini *status predicted cMS2>
F; 1007-1615/Product: nepativirin *status predicted cMS3>
F; 1230-1237/Region: nucleotide-binding motif A (P-loop)
E: 1310/Product: hepatilini *status predicted cMS3>
F; 1310/Production: hepatilini *stat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A.Accession: A39166
A.Molecule type: mRNA
A.Residues: 1-3011 ccDA
A.Residues: 1-3011 ccDA
A.Residues: 1-3011 ccDA
A.Cross-references: GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
A.Cross-references: GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874
B.Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap
J. Gen, Virol. 73, 1131-1141, 1992
A.Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship
A.Reference number: PQ0403
A.Recession: PQ0403
genome polyprotein - hepatitis C virus (strain HCV-1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonst protein NS4s; nonstructural protein NS5)
E;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
C;Accession: A39166; P00403; P00404
R;Choo, Q.L.: Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C. Archo, Q.L.: Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C. Artile: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1663-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match 99.0%; Score 943; DB 1; Length 3011; Best Local Similarity 98.4%; Pred. No. 5.7e-81; Matches 179; Conservative 3; Mismatches 0; Indels 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA A; Residues: 1577-1633 <CHA>
```

ô

Sun Aug 31 15:58:56 2003

E O O

```
A.Note: neither amino acid nor nucleotide sequence is given
C.Superfamily: hepatitis C virus genome polyprotein
C.Sewords: ATP: capsid protein; envelope protein; glycoprotein; hydrolase; nonstruc
C.Sewords: ATP: capsid protein C #status predicted copy.
F:1-115/Product: capsid protein C #status predicted cepw>
F:192-389/Product: major envelope protein H #status predicted cepw>
F:390-729/Product: major envelope protein NSI #status predicted cwsi>
F:300-729/Product: nonstructural protein NSI #status predicted cwsi>
F:1007-1615/Product: nonstructural protein NSI #status predicted cwsi>
F:130-1317/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                      A; Note: host Homo sapiens (man)
C; Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C; Accession: A36814; A41546
R; Inchauspe, G:; Zebede, S:; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A; Description: Genomic structure of the human prototype strain H of hepatitis C viru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013-Product: nonstructural protein NS5 *status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,224
                                                                                                                                    genome polyprotein - hepatilis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein N54a; nonstructural protein N54b; nonstructural protein N55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R;Inchauspe, G.; Zebedee, S.; Lee, Ď.H.; Sugitani, M.; Nasoff, M.; Prince, A.M. Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A;Iiile Genomic structure of the human prototype strain H of hepatitis C virus: A;Reference number: A41546; MUID:92032256; PMID:1658800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A:Molecule type: genomic RNA
A:Residues: 1-3011 <INC>
A:Cross-references: GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:9329738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97.3%;
96.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Species: hepatitis C virus
                                                                                                                                                                                                                                             Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Reference number: A36814
A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ||
|RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     C; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein Na94; nonstructural protein NS40; nonstructural protein NS40; cpecies: hepatitis C virus
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Caccasion: $40770; PC1285
R; Okamoto, H.
Submitted to the EMBL Data Library, March 1992
R; Okamoto, H.
A; Reference number: $40770
A; Molecule type: qenomic RNA
A; Residues: 1-301; OKAA, S: Sugyyama, Y: Y Yotsumoto, S: Tanaka, T: Yoshizawa, H: Tsuda, A; Residues: 1-301; OKAA, S: Sugyyama, Y: Yotsumoto, S: Tanaka, T: Yoshizawa, H: Tsuda, Jpn. J. Exp. Med. 60; 157-177, 1990
A; Rossion: PC1286; MUD: 91013116; PMD: 2170712
A; Reference number: PC1284; MUD: 91013116; PMD: 2170712
A; Reference number: PC1284; MUD: 91013116; PMD: 2170712
A; Reference number: PC1286; Mud: PC1286; Mud
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                    1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ::
ن
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1026 LAPITAYAQQTRGLI.GCIITSI.TGRDKNQVFGEVQIVSTAAQIFLATCINGVCWIVYHGA 1085
                                                                    120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 GDSRGSLLSPRPISYLKGSSGGPLICPAGHAVGIFRAAVCTRCVAKAVDFIPVESLETIM 186
                                                                                                                                                                                                         121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MAPITAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAGTFLATCINGVCWTVYHGA
                                                                                                           GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              F.1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F.1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 #status predicted <N6S>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 939; DB 1;
Pred. No. 1.4e-80;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98.5%;
97.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches 178; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ||
| RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                           RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RS 182
                                                                                                                                                                                                                                                                                                                                          181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                              1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local S
                                                                    61
                                                                                                                          qq
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                     ò
                                                                ŏ
```

```
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                               1085
                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                        180
                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
                                                                                                                                           61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                       121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIM
                                                                                                               1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                        Gaps
                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                      Length 3011;
                                                                     Indels
Score 927; DB 1; Le
Pred. No. 1.9e-79;
                                                                     4; Mismatches
```

```
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,20
                                                                                                                                                              15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          93.1%;
87.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                        Best Local Similarity
Matches 162; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ||
| RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1026
                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       protein NSS pecters. Hepatitis C virus
C;Species: hepatitis C virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
C;Accession: A38465
B;Takamizawa, A.; Mori, C.; Fukc, I.; Manabe, S.; Murakami, S.; Fujita, J.; Cnishi, E.;
J. Virol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from human A;Reference number: A38465; MUID:91140698; PMID:1847440
A;Reference number: A38465; MUID:91140698; PMID:184740
A;Reference number: A38465
A;Molecule type: genomic RNA
A;Residues: 1-3010 c/JAK:
A;Residues: 1-3010 activitis frequence polyprotein
C;Superfamily: hepatitis c virus genome polyprotein
A;Residues: 1-3010 c/JAK:
A;Residues: 1-3010 activitis frequence protein: apredicted c/RS:
F;130-1307/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif A (P-loop
                                                                                                                                                                               A; Cross-references: GB:MB4754
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; hydrolase; nonstructural
C; Superfamily: hepatitis C virus genome protein: divocation c for the constructural
F; 1-115/Product: capsid protein C * status predicted <CPC>
F; 116-191/Product: major envelope protein B * status predicted <MEE>
F; 192-189/Product: major envelope protein RSI * status predicted <NSI>
F; 300-1006/Product: nonstructural protein NSI * status predicted <NSI>
F; 1007-1615/Product: nonstructural protein NSI * status predicted <NSI>
F; 1300-1237/Region: nucleotide-binding motif A (P-loop)
F; 1316-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010,Product: nonstructural protein NS5 #status predicted <NS5>
F:2014-3010,Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4a; nonstructural protein NS4a;
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Reference number: A40244; MJID:92230206; PMID:1314449
A;Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1086 GSKTLAGPKGPITQMYTNVDQDLVGWHAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRR 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTGGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       93.6%; Score 892; DB 1; Length 3010; 90.1%; Pred. No. 4.1e-76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 164; Conservative
                                                                                                                        A; Molecule type: genomic RNA A; Residues: 1-3010 <CHE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
A Macesasion: A July 23 A Molecule type: genomic RNA A, Residues: 1-3010 < KAT>
A, Chossin, S.; Shimotohno, K.
A, Cross-references: GBS 129-223, 1989
A, Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence A, Reference number: PS0086
A, Accession: PS0086
A, Reference number: PS0086
A, Recession: PS0086
A, Recession: PS0086
A, Residues: 2650-2707 < KA2>
A, Reperlamental source: Japanese isolate
C, Commental source: Japanese isolate
A, Residues: 2650-2707 < KA2>
A, Experlamental source: Japanese isolate
C, Comment: The cleavage sites of this polyprotein
C, Comment: The cleavage sites of this polyprotein
C, Superfamily: hepatitis C virus genome polyprotein
C, Superfamily: hepatitis C virus G viru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C.Species: hepatitis C virus
C.Species: hepatitis C virus
C.Species: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C.Species: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C.Species: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C.Species: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jun-2001
C.Species: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jun-1992
A.F. Reference number: A39253; MUID:91088550; PMID:2175903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein · hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61 GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCICGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 MAPITAYAOOTRGLLGCIITSLTGRDKNOVEGEVOIVSTAAOTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1086 GSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                           61 GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GDSRGSLLSPRPISYLRGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF1PVESLETTM
                                                                                                                                                                                                             1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
     Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 887; DB 1; Lengtu vv. Pred. No. 1.2e-75;
93.5%; Score 891; DB 1;
89.0%; Pred. No. 5.1e-76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17; Mismatches
                                                                                                         Mismatches
```

ŭ,

from a single

```
A: Residues: 1.547. TT.5349.621. V', 623-624,'S', 626-652,'DL', 655-761,'T', 763-782 < HOW>
A: Residues: 1.547. TT', 549-621. V', 623-624,'S', 626-652,'DL', 655-761,'T', 763-782 < HOW>
A: Cross-references: EMBL:X61591
A: Note: this sequence is inconsistent with the nucleotide translation
A: Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue:
as Trp, and TTC for residue 771 as Ser
A: Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C: Superfamily: hepatitis C virus genome polyprotein
C: Reywords: ATP: 919-00ptotein; hydrolaeciae cCPC>
F: 115-791-70cduct: capsid protein C *status predicted < CRPA>
F: 106-191/Product: capsid protein WS! *status predicted < NSI>
F: 730-106/Froduct: nonstructural protein NSI *status predicted < NSI>
F: 730-106/Froduct: nonstructural protein NSI *status predicted < NSI>
F: 100-165/Froduct: nonstructural protein NSI *status predicted < NSI>
F: 11317/Region: nucleotide-binding motif B
F: 11317/Region: nucleotide-binding motif B
F: 11312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                    A; Experimental source: isolate Jf from an individual R; Hundami, S. Arch. Virol. 128, 163-169, 1993
A; Ritle: Sequence analysis of putative structural regions of hepatitis C virus isolai A; Reference number: A48332; MUID:93119270; PMID:8380322
A; Rocession: $33570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-32013/Product: nonstructural protein NS4b *status predicted <NNB>
F:2014-3010/Product: nonstructural protein NS5 *status predicted <NS5-
F:2014-3010/Product: nonstructural protein NS5 *status predicted <NS5-
F:196,209,234,250,305,411,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                     C:Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 23-Mar-2001 C:Date: 19-May-2000 *sequence_revision 199059 C:Accession: $180030: $35570; A48329 S: $180029 C:Ancession: $18003 C:Ansashi, U.; Kobayashi, K.; Murakami, S. submitted to the EMBL Data Library, September 1991 C:Ancestion: A whole genome of hepatitis C virus cDNA was isolated from a A;Reference number: $18028 A;Accession: $18030
                                                                                                                                                                                                                                                                 A;Molecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross-references: EMEL:X61596; NID:959478; PIDN:CAA43793.1; PID:959479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
hes 158; Conserv
isolate JK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3C5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A.Cross.references: 13010 v.mm.
A.Gross.references: GB.011168, GB.001171; NID:9221612; PIDN:BAA01943.1; PID:9221613
A.Experimental source: HCV-JT
A.Note: sequence extracted from NCB1 backbone (NCBIN.106206, NCBIP.106207)
C.Superiaminy: hepatitis C virus genome polyprotein
C.Superiaming Protein Protein C *status predicted CEPN>
F.2-115/Product: capsid protein M *status predicted CEPN>
F.190-239/Product: major envelope protein E *status predicted CNSI>
F.190-239/Product: nonstructural protein NS1 *status predicted CNS2>
F.1007-1615/Product: nonstructural protein NS2 *status predicted CNS2>
F.1230-1237/Region: nucleotide-binding motif A (P-loop)
F.1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (isolate JKl)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a: nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Virus Res. 23, 39-53, 1992
Virus Res. 23, 39-53, 1992
A;IIIIe: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: 8, Reference number: A45573; MUID: 92295714; PMID: 1318627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11 46 GDGRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVESMETIM 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1086 GSKTLAGPKGPITQMYTNVDQDLVGWHAPPGARSLTPCTCGSSDLYIJVTRHADVIPVRRR 1145
                                                                                                                                    GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                    121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
                            .086 GSKTLAGPKGPITQMYTNVDQDLVGWPAPPGARSMTPCTCGSSDLYLVTRHADVVPVRRR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <N4A>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F.1616-1862/Product: nonstructural protein NS4a *status predicted <N4A: F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B: F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 2.3e-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92.8%; Score 884;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Molecule type: DNA
A;Residues: 1-3010 <TAN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
Matches 161; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RS 1207
                                                                                                                                                                                                                                                                                                  RS 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RS 182
                                                                                                                                                                                                                                     RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                     1206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ထ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S18030
                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       \delta
                                                                                                                                                                 g
                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                          g
```

```
genome polyprotein - hepatitis C virus (isolate EUK1480)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru protein NS4s; nonstructural protein NS5)
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Date: 19-May-2000 *text_change 19-Jan-2001
C; Accession: JC5520
B; Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem Biophys. Res. Commun. 236, 44-49, 1997
A; Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomine A; Reference number: JC5520; MUID: 9736559; PMID: 9223423
                                                                                                                                                                                                                  1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                                                                                                                                                 121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                    1086 GSKILAGPKGPINOMYTNVDQDLVGWQAPSGAASLTPCTYGSSDLYLVTRHADVIPVRRR
                                                                                                                                                  1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                          GTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
       Length 3010;
                                                                       Indels
90.7%; Score 864; DB 1; E. 86.8%; Pred. No. 1.9e-73; ive 15; Mismatches 9;
```

C:Species: hepatitis C virus

```
Octobrance polyprotein - hepatitis C virus (isolate HC-J6)

Wicontains: capsid protein C; envelope protein M: hepacivirin (EC 3.4.21.98) (nonst protein NS4s; nonstructural protein NS4s; nonstructural protein NS4b; nonstructural protein NS5

C;Species: hepatitis C virus

C;Species: hepatitis C virus isolated from C;Accession: J03303

A;Mclecule vipe: genomic RNA

A;Ritle: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from A;Accession: J03303

A;Mclecule type: genomic RNA

A;Residues: J033 000844; NID:9204440; PMID:1658196

A;Mclecule type: genomic RNA

A;Residues: J033 000844; NID:9221650; PIDN:BAA00792.1; PID:9221651

A;Reperimental source: isolate HG-J6 from a Japanese individual

C;Superfamily: hepatitis C virus genome polyprotein;

C;Superfamily: hepatitis C virus genome polyprotein

C;Superfamily: hepatitis C virus genome

C;Superfamily: hepatitis C virus genome

C;Superfamily: hepatitis C virus
A;Wolecule type: mRNA
A;Residues: 2678-2799 cKAI>
A;Residues: 2678-2799 cKAI>
A;Residues: 2678-2799 cKAI>
A;Cess.references: GB.D10562; GB.D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
A;Cross.references: GB.D10562; GB.D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruct
F;116-191/Product: envelope protein M *status predicted <EPM>F:192-389/Product: major envelope protein B *status predicted <NS1>F:390-733/Product: nonstructural protein NS1 *status predicted <NS2>F:1011-1619/Product: nonstructural protein NS1 *status predicted <NS2>F:1011-1619/Product: hepacivirin *status predicted <NS3>F:1234-1241/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                    F:1320-1323/Region: DEXH motif
F:1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>
F:1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
F:2018-3033/Product: nonstructural protein NS5 *status predicted <NS5>
F:196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,20:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 74.9%: Score 714; DB 1; Length 3033; 70.9%; Pred. No. 3.5e-59; 1ve 26; Mismatches 27; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 70.9%;
Matches 129; Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |:
|210 RT 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             301303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                            A; Molecule type: genomic RNA
A; Residues: 1-3033 COKA>
A; Cross-references: GB:DD1929; GB:DD1921; NID:g221608; PIDN:BAA01761.1; PID:g221609
A; Cross-references: GB:DD198; GB:DD1921; NID:g221609; PIDN:BAA01761.1; PID:g221609
J. Gen. Virol. 73, 1131-1141, 1992
A; Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to GA; Reference number: PQ0393; MUID:92268871; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Experimental source: isolate E-bl2
R; Kato, N.; Octsuyama, Y.; Ochkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohnd
Blochem. Blophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C;Species: hepatitis C vitus
C;Species: hepatitis C vitus
C;Species: hepatitis C vitus
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A40250; p00397; p00559
R;Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.;
Virology 188, 331-341, 1992
A;Title: Full-length sequence of a hepatitis C virus gcnome having poor homology to repd
A;Reference number: A40250; MUID:92230232; PMID:1314459
A;Accession: A40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (strain iiC-38)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTKHADVIPVRRR 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2210-2249/Region: interferon sensitivity determining *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84.4%; Score 804; DB 1; 79.1%; Pred. No. 9.5e-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Molecule type: genomic RNA
A;Residues: 2678-2754 <CHA>
A;Cross-references: DDBJ:D10134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 144; Conservative
                                      Residues: 1-3014 <CHA>
Cross-references: GB:Y13184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
          Molecule type: mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RS 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
```

```
C; Species: Homo sapiens (man)
                                                                                                          Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Status: preliminary A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: H83144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   109
                                                                                                                                                   Matches
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Polyprotein - douroucouli hepatitis GB virus A C:Species: dourous C:Species: J.C.; Dosai. S.M.: Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Genomic analysis of two GB virus A variants isolated from captive monkeys. A;Reference number: 216486; MCID:98120818; PMID:9460920 A;Reference number: 216486; MCID:98120818; PMID:9460920 A;Status: translated from GB/EMBL/DDBJ A;Molecule type: mRNA A;Status: translated from GB/EMBL/DDBJ A;Molecule type: mRNA A;Cross-references: EMBL:AF023425; NID:92828599; PIDN:AAC40502.1; PID:92828600 C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              R.Erker, J.C.: Desai, S.M.: Leary, T.P.: Chalmers, M.L.: Montes, C.C.: Mushahwar, I.K. J. Gen. Virol. 79, 41-45, 1998
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A;Reference number: 216486; MUID:98120818; PMID:9460920
A;Reference number: T08839
                                                                                                                                                                                            • •
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1038 ARTLAGPVGFVNCRWWSPSDDVAVYPLPSGASCLEPCKCGTQSVWCIRN--DGALCHGRL 1095
                                                                                                            .;
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62 TRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRG 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 DSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGV-----AKAVDFIPVE 174
                                                                                                                                                                                                                                                      GDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTM 180
                                                                                        09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 APITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein - marmoset hepatitis GB virus A
C;Species: marmoset hepatitis GB virus A
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000
C;Accession: 100839
                                                                                        1 MAPITAYAQQTRGLLGCITTSLTGRDKNQVEGEVQIVSTAAQFFLATCINGVCWTVYHGA
                                                                                                                                                                        GTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRR
                                                   Gaps
                                               ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3065;
      Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28.1%; Score 267.5; DB 2; 33.1%; Pred. No. le-16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               78;
    74.7%; Score 712; DB 1; L. 69.8%; Pred. No. 5.4e-59; Live 29; Mismatches 26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 32; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
Query Match
Best Local Similarity 69.8%;
Marches 127; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 28.1%
Best Local Similarity 33.1%
Matches 60; Conservative
                                                                                                                                                                                                                                                                                                                                                                ||
RS 1211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             S 1155
                                                                                                                                                                                                                                                                                                                                      181 RS 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      175 S 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
T08839
                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                             a
                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                    á
                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                           a
```

```
Cispecies: Pseudomonas aeruginosa (str. Cispecies: Pseudomonas aeruginosa PAOI, administic p. A. Recence (str. Cispecies: Pseudomonas aeruginosa PAOI, an opportunistic p. A. Recence (str. Cispecies: Paulos) (str. Cispecies: Paulos) (str. Cispecies: Pseudomonas aeruginosa PAOI, an opportunistic p. A. Recence (str. Cispecies: Pseudomonas aeruginosa PAOI, an opportunistic p. A. Recence (str. Cispecies: Pseudomonas aeruginosa PAOI) (str. Cispecies: Pseudomonasa aeruginosa PaOI) (str. Cispecies: Pseudomonasa aeruginosa PaOI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A:Residues: 1-09 <STO>
A:Cross-references: GB:AE004818; GB:AE004091; NID:g9950200; PIDN:AAG07406.1; GSPDB:GI
A:Experimental source: strain PA01
C:Genetics:
C:Genetics:
C:Superfamily: dedF protein
                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     970 APVIVIH-HHGKGFFGVVKISMIGNDETEHVGNVVVLGISIIRSMGICVNGVMYIIYHGSN 1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       67 SPKGP-----CTCGSSDL 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 LPAKPQAMQAFLTEYCGAAAGQIRVFGQND-----WMAPPASGSSAPNAMVICPCSTGTL 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----YLVIRHADVIPVRRRGDSRGSLLSPR--PIS-----YLKGSSGGPLLCPA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 TRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG 121
A;Cross-references: EMBL:AF023424; NID:92828597; PIDN:AAC40501.1; PID:92828598 C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8 AQQTKGLLGCIITSLTGRDKNQVEGEVQ-IVSTAAQTFLATCINGVCWTVYHGAGTRTIA 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2 APITAYAQOTRGLLGCIITSLTGRDKNOVEGEVQIVSTAAQTFLATCINGVCWTVYHGAG 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DSRGSLLS-----PRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                76; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                         33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 9.1%; Score 87; DB 2; Length 209; Best Local Similarity 26.2%; Pred. No. 0.69; Matches 55; Conservative 18; Mismatches 51; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     17 AQYGLRLLDCLV-----OEEREVHFLISKAAQLVMAT-----
                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                        ch 26.8%; Score 255.5; DB 2; Il Similarity 30.1%; Pred. No. 1.4e-15; 59; Conservative 36; Mismatches 68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            149 GHAVGIFRAAVCTRGVAKAVDFIPVESLET 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       165 --APGFYHQ---PQSVEDLVDFVVARILNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   angio-associated migratory cell protein - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1135 YTKPWETLPREAITHT 1150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               164 VAKAVDFIPVESLETT 179
```

```
Cyaccession: 199383
R;Beckner, M.E.; Krutzsch, H.C.; Stracke, M.L.; Williams, S.T.; Gallardo, J.A.; Liotta, Cancer Res. 55, 2140-2149, 1995
A;Title: Identification of a new immunoglobulin superfamily protein expressed in blood wa; Reference number: 139383; MUID:95262124; PMID:7743515
A;Reference number: 139383; MUID:95262124; PMID:7743515
A;Reference number: 139383; MUID:95262124; PMID:7743515
A;Reference number: 139383; MUID:97262124; PMID:7743515
A;Reference number: 139383; MUID:9870802; PIDN:AA68889.1; PID:9870803
A;Cross-references: GB:M95627; NID:9870802; PIDN:AA68889.1; PID:9870803
C;Genetios:
A;Genetios: A;Genetios: A;Genetios: A;Genetios: A;Genetios: A;Genetios: A;Genetios: A;Genetios: Ulaq32.1-14q32.1
C;Superfamily: unassigned Wp repeat proteins: WD repeat homology wwb:>F;148-181/Domain: WD repeat homology wwb:>F;414-447/Domain: WD repeat homology wwb:>F;414-447/Domain: WD repeat homology wwb:>F;414-447/Domain: WD repeat homology wwb:>F;414-447/Domain: WD repeat homology wwb:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            109 TRHADVIPVRRR---GDSRGS-----LLSPRPISYLKGSSG--GPLLCPA----- 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 MTVYHGAGTRITASPKGPVIQMYTNVDKDLVGWPAPQGSRSL.----TPCTCGSSDLYLV 108
C;Date: 06-Sep-1996 *sequence_revision 06-Sep-1996 *text_change 21-J:ii-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
8.6%; Score 82; DB 2; Longth 452;
Best Local Similarity 25.3%; Pred. No. 5;
Matches 42; Conservative 13; Mismatches 47; Indels 64; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 149 ------GHAVGIFR----AAVCTRGVAKAVDFIPVESI, 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296 GSVDCQAKLVSATIGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              do
```

Search completed: August 30, 2003, 19:20:24 Job time : 17.9789 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

```
August 30, 2003, 19:20:43; Search time 1764.86 Seconds (without alignments) 2506.388 Million cell updates/sec
                                                                                                                           1 MAPITAYAQQTRGLLGCIIT......GVAKAVDFIPVESLETIMRS 182
- nucleic search, using frame_plus_p2n model
                                                                                                                                                                            0.5
                                                                                                                                                          BLOSUM62
Xgapop 10.0 , Xgapext (
Ygapop 10.0 , Ygapext (
Fgapop 6.0 , Fgapext (
Delop 6.0 , Delext (
                                                                                          US-09-965-594-1
953
                                                                                            Title:
Perfect score:
Sequence:
                                                                                                                                                               Scoring table:
OM protein
                              Run on:
```

Total number of hits satisfying chosen parameters:

22781392 seqs, 12152238056 residues

Searched:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-MODEL-frame+_p2n.model -DEV=xlp
-MODEL-frame+_p2n.model -DEV=xlp
-C-gnz_1/USPTO_spool/US09965594/runat_29082003_151919_28322/app_query.fasta_1.2872
-De-Cgnz_1/USPTO_spool/US09965594/runat_29082003_151919_28322/app_query.fasta_1.2872
-DE-EST_CPHT-fastap -SUFFIX=rst -MINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END=-1 -MATRIX=blosum62 -TRANS=buman49 cd1 -LIST-45
-UNITS-bits -START-1 -END=-1 -MATRIX=DIOSUM62 -TRANS=buman49 cd1 -LIST-45
-OUTPHT-pto -NORM-ext -HEAFSIZE=500 -MINELN-0 -MAXLEN-2000000000
-USBE-US0965594 @CGN_1 1_12630_@runat_2908303_151919_28322 -NCPU-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORES-0 -MAIT -DSPBLOK=100 -LONGLOG
-DEV_IMBOUT-120 -WARN_TIMBOUT-30 -IHREADS-1 -KGAPOP-10 -KGAPEXT-0.5 -FGAPOP-6
-FGAPEXT=7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP=6 -DELEXT=7

Database :

em_gss_phg:* em_gss_vrl:* qb_gssl:* em_gss_hum:* em_gss_inv:* em_gss_pln:* em_gss_vrt:* em_gss_fun:* em_gss_mam:• em_gss_mus:* em_gss_pro:* em_gss_rod:* em_estfun:* em_estom:* em_estba:*
em_estin:*
em_estin:*
em_estov:*
em_estov:* em_estro: em_htc:* gb_est1:* gb_htc:* gb_cst3:* gb_cst4:* gb_est5:* EST:*

gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

v;	
ES.	
\vdash	
3	
4	
3	
ç	
=	
×	
S)	

Homo sapiens Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 984) BF304699.1 GI:11251586 mRNA sequence. BF304699 EST. ORGANISM ACCESSION REFERENCE VERSION KEYWORDS SOURCE

a

```
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               574
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                     AUTHORS
TITLE
JOURNAL
COMMENT
                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                            1. 984
//organism="Remo sapiens"
//organism="Remo"
//organism="Remo"
//db_xref="taxon:9606"
//db_xref="taxon:9606"
//db_nost="lyhe"=Trhabdomyosarcoma"
//db_host="lyhe"=Trhabdomyosarcoma"
//db_host="lyhe"=Trhabd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AGENCOORPE 8417538 Lupski_sympathetic_trunk Homo sapiens cDNA clone INAGE:6192708 5', mRNA sequence.
BQ892487 GI:22284501
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     105 LeuTyrLcuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArg 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               125 GlySerLeuLeuSerProArgPro-------IleSerTyrLeuLysGlySer 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               140 SerGlyGlyProLeuLcuCysProAlaGlyHisAlaValGlyIlcPheArgAlaAlaVal 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    493 CGGCACGCTCCGGTCAGGTGCAGC-----------TTCCAGCGCCCGGGG 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TrpProAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          598 ------ACCAGGCACGCAGCGAACATACATGCAAGGAGACGTGGT---TCCCGC 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    553 GGCCCCTCTTGTGGGGAAGACCTCGATGGTGTCCAAGCTCGCGCTGCTGTACTGGAAGT 494
                                                                                                             Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCH1005 row; g column: 13
High quality sequence stop: 646.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       984
33
5
22
5
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                           Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-1 (1-182) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.34
105.00
45.248
39.298
11.128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         160 CysThrArgGly 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TGCGCCCGAGGA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HOMO
HOMO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 2
BQ892487
                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

.

```
/^clone_lib="Lupski_sympathetic_trunk"
//ote="Vector: pCWV-SPORT6 (Life Technologies); Site_l:
Noti="Vector: pCWV-SPORT6 (Life Technologies); Site_l:
Notis Site_2: Sall; cDNA made by oligo-dr priming.
Directionally cloned using the following adaptors:
5. TGGACCACGCGCGTGCG 3' and
5. GACTAGTTCTAGATCGCGAGGGCGCCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          484 TGGGATCCATTTTTAATAAGGGTGCTCTGTTAATCATGGCGCCCACGGGCCCGCTGATA 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------CysGlySerSerAsp 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           105 LeuTyr-LeuValThr------ ArgHisAlaAspVallleProValArg---- 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -------ArgArgGlyAspSerArgGlySerLeuLeu-- 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----SerProArgProIleSerTyrLeuLys-------GlySerSerGl 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIle 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74 GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArg 93
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1199)
                                                                                                                                                                                              Couract. NODE: Strausberg, Fir.U.
Fmail: cgapbs-rémail.nih.gov
Tissue Procurement: Dr. James R. Lupski
conN Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://amage.llnl.gov
Plate: LLAM13595 row: c column: 13
High quality sequence start: 57
High quality sequence stop: 394.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ---TGTGCTGCCTGCACACACCCCCATGACCGATGTTGGGGCTTATGTGGAAACGGCGAG
                                                                                     NIH'MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                          þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 41
17
53
64
64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    211 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1. .1199
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="!NAGE:6192708"
/sex="male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
                                                                                                                                                                              Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-1 (1-182) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 343 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerLeuThrProCysThr ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Technologies.
362 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103.50
37.428
26.458
10.868
```

3

296

96

16

26

114

8

οý 5 ò

```
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BM374064 53-JUL-2002 EBma03_SQ003_I16_R maternal, 8 DPA, no treatment, cv Optic, EBma03 Hordeum vulgare subsp. vulgare cDNA clone EBma03_SQ003_I16 5', mRNA
   189 GGAGGGCG-------GCATGTGTTCTTCTGCTTCTGGTGGACCAGCGGGGA 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          297 TGGCGGACA-------TGGAGACGGCCACCCGGCTCTGCTACCTCCTCA 338
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 lileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSe 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 399 CCTACCG------CCACTACCGGCTCTACTAGACCCGCGTCGTTGAC 440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | (Pases 1 to 583)
| Hedley, P., Liu, H., Caldwell, D., McCallum, N., Mudie, S., Cardle, L.,
| Ramsay, L., Machray, G., Marshall, D. F.M. and Waugh, R.
| Development of Barley Transcriptome Resources
                                                                                        ---GlyGluValGlnIleValSer
                                                                                                                                                                                                                                                                                                                                                               77 ThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            97 ProCysThrCys-----GlySerSerAsp-LeuTyrLeuValThrArgHisAlaAspVa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                                                                                                                                                                                                          39 ThrAlaALaGlnThrPheLeuAlaThrCyslleAsnGlyValCysTrpThrVal-----
                                                                                                                                              130 CATCAAATACC-TCCGGCGAGAGTTTCAGCGATATTTCTGGAGGAGCTTGGCGCGGAAGGA
                                                                                                                                                                                                                                                                                                                            57 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /clone_lib-"maternal, 8 DPA, no treatment, cv Optic, EBma03"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      On Jan 10.002 this sequence version replaced gi:18117454. On Jan 10.02 this sequence version replaced gi:18117454. Contact: Waugh R, Marshall DF Genome Dynamics/Computational Biology Scottish Crop Research Institute
Invergowrie, Dundee, DD2 5DA, Scotland, UK
Tel: 00 44 1382 562731
Fax: 00 44 1382 562426
Email: estéscrisariac.uk
All sequence has a pired quality score of 20 or over Seq primer: M13 reverse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Hordeum vulgare subsp. vulgare"
nou_type="mRNM"
/cultivar="Optic"
/db_xref="taxon:112509"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /clone="EBma03_SQ003_I16"
/tissue_type="maternal"
/dev_stage="8 DPA"
/lab_host="DH108"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BM374064.2 GI:21941288
                                                                                     28 AsnGlnValGlu---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       134 rTyrLeu 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               441 GCTGATT 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FFATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BM374064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C.Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                               õ
                                                                                                                                    g
                                                                                                                                                                                       6
                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                    \overset{\sim}{o}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /dobe_lbb_H2.
//clone_lbb_H2.

                                                                                                                                                                                                                                                                                                                                                                                        CA023748 13-OCT-2002
LAGYELIZ HZ Hordeum vulyare subsp. vulqare CDNA clone HZ47El7
5-PRIME, mRNA sequence.
CA023748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota, Viridiplantae: Štreptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
AAAACCCCACCTICGGCCCCACCGATGCGCTACCCTTTACAAGCCACCGCCGGG 810
                                                                                                                                              811 CCCCCCCTAACATCTCCTACCCTGCGGCGCGGGGGGGGGAGACGTGGCGCATACGGGC 870
                                                                                     yGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysTh 16:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Wobus, U.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Grop Plant Research (IPK)
Corrensetr 3, 06466, Gatersloben, Germany
Tel: 039482-5522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /tissue_type="pericarp"
/dev_stage="0-7 DAP (days after pollination)"
/lab_host="XLIO-Gold"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism-"Hordeum vulgare subsp. vulgare"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Potokina, E. and
                                                                                                                                                                                                          161 rArgGlyValAlaLysAJaValAspPhelleProValGluSer 175
                                                                                                                                                                                                                                                                      871 TCAGGGCGTTTTAAAGCCCCCCGGCCTTCGCCGCGGGGGAAGCA 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              515
44
16
16
53
31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Email: stein@ipk-gatersleben.de
Insert Length: 515 Std Error: 0.00
Platc: 47 row: E column: 17
Seg primer: M13rev.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Radchuk,V., Zhang,H., Weschke,W.,
Barley ESTB from developing seeds
Unpublished
Contact: Stein Nils
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /db_xref="GABI:277158"
/db_xref="taxon:112509"
/clone="HZ47E17"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hordeum vulgare subsp. vulgare
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hordeum vulgare subsp. vulgare
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-1 (1-182) x CA023748 (1-515)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type="mRNA"
/cultivar="barke"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CA023748.1 GI:24301322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Triticeae; Hordeum.
1 (bases 1 to 515)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.6
101.00
41.968
30.778
10.608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Fax: 039482-5595
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                     141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
                                                                                                                                                                                                                                                                                                                                  RESULT 3
CA023748
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
```

```
140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BJ02412
                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  373
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                           ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
BJ024121
                                                FEATURES
                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
Non-normalised library, directionally cloned into pspoRTI. Derived from maternal Lisaue disserted from developing grains (8 days post anthesis) in glasshouse grown barley plants. Developed as part of the barley transcriptome resources of BBSMC/SERRAD funded cereal IGF (Investigating Gene Function) profect.

a 176 g 141 t.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BJ001625
BJ001625 MF01SSA CDNA OFYZIAS latipes CDNA clone MF01SSA025C02 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      279
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinoptorygii; Neoptorygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProCysThrCys-----GlySerSerAsp-LouTyrLeuValThrArgHisAlaAspVa 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 lileProvalArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSe 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 (bases 1 to 615)
Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                                                                           71 CATCAAATACC-TCCGGCGAGAGTTTCAGCGATATITCTGGAGAGCTTGGCGGGAAGGA 129
                                                                                                                                                                                                                                                                                                                                                                                                                                   130 GGAGCGGCG------GCATGGTGTTCTTCTGCTTCCTGGTGGACCAGCGGCGGA 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96
                                                                                                                                                                                                                                                                                                      11 GCACAGGGGGGGGGGGGCTAGCTGTATTGCACTGTCCACCACCACGAGGCACCAT 70
                                                                                                                                                                                                                                                                                       AlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThrSerLeuThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   57 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      238 TGGCGGACA-------TGGAGACGGCCACCCGGCTCTGCTACCTCCTCA
                                                                                                                                                                                                                                                                                                                                                -----GlyGluValGlnIleValSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            178 AGTGCGGAGCAGCAAGCCGGCGGCGGCATCTGCTCGCGGTGCGGCGCTGCGCCAGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          77 ThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThr
                                                                                                                                                                                                                                                                                                                                                                                                       39 ThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrVal----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Beloniformes; Adrianichthyidae; Oryzinae; Oryzias.
                                                                                                                                                     583
16
53
53
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Oryzias latipes (Japanese medaka)
                                                                                                                                                                                                                                                          (1-583)
                                                                                                                                                                                                                                                                                                                                                AsnGlnValGlu------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BJ001625.1 GI:17364516
                                                                                                                                                                                                                                                          US-09-965-594-1 (1-182) x BM374064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Contact: Tadasu Shin-i
                                                                                                                                                  6.6
101.00
41.96%
30.77%
10.60%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Oryzias latipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GCTGATT 388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        rTyrLeu 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unpublished
                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BJ001625
                                                                                                                                                                                 Percent Similarity:
                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EST
                                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97
                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
BJ001625/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                      Pred. No.:
                                                                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
BJ024121
BJ024121 MF01SSA CDNA Oryzias latipes CDNA clone MF01SSA143D12 3',
MRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
extinopterygii; Neopterygii; Teleostei; Buteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryzinnae; Oryzias.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124 ArgGlySerLeuLeuSerProArg-----ProlleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------ACTCGTGGAGGACGAAGAGCGTCACCCGGAGCTGTAGG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   223 AGAGGCCGGACAGCAGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         47 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             67 SerProLys-------GlyProValileGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   283 TCTTCCTCACCTGACCTTCCACATCCAGGGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg
                                                                                                                                                                                                                                                                                                                                                                                            615
42
9
50
50
                                                                                                                      / ...yuram - ...yuras taripes
/strain - Hd-rR"
/strain - Hd-rR"
/db_xref - Ttaxon: 8090"
/clone - MF01SSA025C02"
/sex - mixture of female and male"
/tissue_type - "whole embryo"
/dev_stage - seymentation stage 20 - /clone - 1b - MF01SSA cDNR"
                                                                                                                                                                                                                                                                 ,
                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlyGlyProLeuLeuCysProAlaGlyHisAla 151
                                                                                                             /organism-"Oryzias latipes"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Oryzias latipes (Japanese medaka)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-1 (1-182) x BJ001625 (1-615)
BJ024121.1 GI:17377389
                                                                                                                                                                                                                                                                                                                                                                                            11
99.00
33.77%
27.81%
10.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Oryzias latipes
                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
```

S

```
------CysThrCysGlySerSerAspLeuTyrLeuValThrArg------ 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     111 .....HisAlaAspValIleProValArgArgArgGlyAspSer 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124 ArgGlySerLeuLeuSerProArg-----ProlleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 530 AGAGGCCGGACAGGCAGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA 589
                             Oryzias latipes (Japanese medaka)
Oryzias latipes
Oryzias latipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Eureleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryzinae; Oryzias;
1 (bases 1 to 754)
Medaka EST Project in Takeda's lab
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        27 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        47 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             67 SerProLys------GlyProValIleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     170 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     82 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - 25
                                                                                                                                                                                                                                                                                                                                                               1...754
/organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="Mr01SSA025C02"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="seqmentation stage 20 -
/clone_llb="Mr01SSA CDNA"
| a 181 c 181 g 198 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              754
42
9
50
50
                                                                                                                                                                                                                                                                       1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        141 GlyGlyProLeuLeuCysProAlaGlyHisAla 151
                                                                                                                                                                                                                    Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-1 (1-182) x BJ016176 (1-754)
                                                                                                                                                                                                                                                                                                                                Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
BJ016176.1 GI:17376695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              14.4
99.00
33.77%
27.81%
10.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            302
                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ..
O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
                                                      ORGANISM
                                                                                                                                                              AUTHORS
TITLE
JOURNAL
COMMENT
VERSION
KEYWORDS
SOURCE
                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BJ016176 TF BF015SA CDNA Oryzias latipes CDNA clone MF01SSA025C02 3', BJ016176 MF015SA025C02 3', BJ016176 MF015SA025C02 3', BJ016176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ArgGlySerLeuLeuSerProArg-----ProTleSerTyrLeuLysGlySerSer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  530 AGAGGCCGCACAGCCAGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA 589
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             302 -------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGCCA 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       242 AAAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
         Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82 AspleuValGlyTrpProAlaProGlnGlySerArgSerLcuThrPro------ 97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    27 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          47 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------ACTCGTGGAGGACGAAGAGCGTCACCCGGGAGCTGTAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---HisalaAspVallleProValArgArgArgGlyAspSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              470 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67 SerProLys ------GlyProVallleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                        - 254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  643
90
50
50
                                                                                                                                                                                                                                                                                        /db.xref."taxon:8090"
/clone="MF01SSA143D12"
/sex="mixture of female and male"
/tissue_typer"whole embryo"
/dev_stage="segmentation stage 20 -
/clone_lib="MF01SSA cDNA"
a 148 c 148 q 176 t
                                                                      Contact: Tadasu Shin-i
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata: Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6856
Email: tshini@qenes.nig.ac.jp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141 GlyGlyProLeuLeuCysProAlaGlyHisAla 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism-"Oryzias latipes"
/mol_type-"mRNA"
/strain-"Hd-rR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-1 (1-182) x BJ024121 (1-643)
   (bases 1 to 643)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    111.7
99.00
33.77%
27.81%
10.39%
                                                                                                                                                                                                                    .643
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                              171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           111 ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 7
BJ016176
                                       TITLE
JOURNAL
   REFERENCE
                   AUTHORS
                                                                                                                                                                                                        FEATURES
                                                                            COMMENT
```

a à

g

g οy

Š

ð a q ò q ò

9

```
/organism="Mus musculus"
//organism="Mus musculus"
//db_Lype="mxRA"
//db_Lxref="taxon:10090"
//db_host="mil0B (TI-phage-resistant)"
//clone_lib="NIH_MGC_17"
//clone_lib="NIH"
//clone
                                                                                                                                                                                                                                                                    EST 29-APR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     332 GGAIGTCICATIGATIAIGIGCCTCAAGIGIGCCACCIGCTCACCCACAAGGIICAGIAI 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             392 GIGATICAAGGCIAICCCCAAAACAAGAGAGIACAICGGCAGCITITCCIICACICAIT 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E (Amanalla) Distribution, Notering, Schillegierin, Fulluae; Fulliae; Fulliae; Full (Bases 1 to 1031)

S NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strausberg, Ph.D.
Email: Gapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CONA Library Preparation: Michael Brownstein Laboratory
CDNA Library Preparation: Michael Brownstein (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NCCMIO7 row: b column: 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29 GlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCys 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16 GlyCysIleIleThr-------SerLeuThrGlyArgAspLysAsn 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          49 IleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
                                                                                                                                                                                                                                                                CB950999 1031 bp mRNA linear EST
AGENCOURT_13445496 NIH_MGC_177 Mus musculus cDNA clone
IMAGE:30316162 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1031
47
16
59
35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      =
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                High quality sequence stop: 333.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-1 (1-182) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                CB950999.1 GI:30205777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24.3
98.50
40.38%
30.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
                                        166 ys 166
                                                                                                            417 AA 416
                                                                                                                                                                                                                                                                                                                                                                                CB950999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                                                                                                                                                                LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                        RESULT 9
CBS50999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1. 961
//organism="Homo saplens".
//organism="Homo saplens".
//mol_type="mmna".
//db_xref="Taxon:9606"
//db_bost="Thabdomyosarcoma"
//db_host="Thabdomyosarcoma"
//db_host="Dh10B (plage-fesistant)"
//db_host="Dh
                                                                                                                                                    BF203316 961 bp mRNA linear EST 06-NOV-2000 601865914F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4098578 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126 rLeuLeuSer-ProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuC 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             531 CATGCTGTCCTCCTCGTCCGCGGGGAGTAGGCAAGGGTCTGGAGGAGGAGGG-----T 478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer-AspLeuTy 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaL 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Oniqualization
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: ATCS
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM965 row: I column: 03
High quality sequence stop: 637.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 (bases 1 to 961)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     961
32
32
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative:
Mismatches:
590 GGAGGACCGACTCGCIGCAGAGCCTCTGCTGCA 622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-1 (1-182) x BF203316 (1-961)
                                                                                                                                                                                                                                                                                                BF203316.1 GI:11096902
                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22.1
98.50
52.44%
39.02%
10.34%
                                                                                                                                                                                                                            mRNA sequence.
BF203316
                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                        DEFINITION
                                                                                                        BF203316/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                ACCESSION
                                                                            RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
```

ð g õ qq à Q ç g

10 10 10 10 10 10 10 10	Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C., Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D., Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P., Marchionni,L., Mashima,J., Mazzarelli,J., Monbaerts,P., Nordone,P., Ring,B., Ringwald,M., Seya,T., Shamanto,N., Sasaki,H., Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Sizuki,H., Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L., Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S. and Hayashlazki,Y. Nature 409 (6821), 685-690 (2001) 2108566 11217851 Group Phase i ill Team. Group Phase i ill Team.	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs L Ature 420, 563-573 (2002) E (bases 1 to 1141) S Adachi, J., Alzawa, K., Akimura, T., Arakawa, T., Bono, H., Carnincl, P., Fukuda, S., Furuno, M., Hangaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horih, F., Inchara, K., Stalia, Y., Kondo, S., Konno, H., Kasukawa, T., Kasukawa, T., Katoh, H., Kawai, J., Kojim, Y., Kondo, S., Konno, H., Kouda, M., Nowlas, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Salto, R., Saltoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Saltoh, H., Sakai, T., Tanaka, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y., Toya, T., Yasunishi, A.,	Direct Submission Direct Submission Direct Submission Direct Submission Direct Submission Direct Submission Division and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Vokohama Institute; 1-7-22 Suchlro-cho, Tsurumi'ku, Yokohama, Ranagawa 230-0045, Japan (E-mail:genome-resegsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-922, Fax:81-45-503-9216, CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/. URL:http://fentome.gsc.riken.go.jp/. URL:http://fentome.gsc.riken.go.jp/.		// Ance-"unimated protein product; purative weakly similar to zinc finger protein (fragment) [Mus musculus] (PIR]148722, evidence: FASTY, 50.7%ID, 57.6%length, match-601) // S. (%length, match-601) // Codon, start-3 // Protein_id-"BAC37940.1* // Ab_xref-"G1:634860.1* // Ab_xref-"
199 TTGGGCACACTGGTCGCACAT	TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	TITLE JOURNAL REFERENCE AUTHORS	🐔	FEATURES SOURCE	polyA
THE MED E EDWERTEED W EREZO	LysGlyProvalileGlnMetTyrThrAsnValAspLysAspLeuVaiGlyTrpProAla [1] [1] [1] [1] AAAGGGCTTTCACAAAAACACTTAACCTCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	AK080545 AK0	Mus musculus Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; But Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu. Larninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 9279253 10349636 Zarninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shiba Ich,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashi Normalization and subtraction of cap-trapper-sclected operowe Res. 10 (10), 1617-1630 (2000)	11042159 3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H. Sumi,N., Ishii,Y., Nakamira,S., Hazama,M., Nishine,T., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Inegami,T., Ke Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., PYoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,Sokazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayash, RIKRN intograted sequence analysis (RISA) system-384-1864000 Res. 10 (11), 1757-1771 (2000)	

polyA_sitc 1141 Note="putative" Anote="putative" ORIGIN	COMMENT On Dec 19, Contact: Wi Clemson Uni Clemson Uni
Alignment Scores: 27.7 Length: 1141 Score: 98.50 Matches: 51 Percent Similarity: 34.83% Conservative: 11 Best Local Sinilarity: 28.65% Mismatches: 59 Query Match: 10.34% Indels: 57 DB:	100 Jordan Tel: 864 65 Fax: 864 65 Fax: 864 65 Email: rwin Total hq ba Seq primer: High quaer:
US-09-965-594-1 (1-182) x AK08U545 (1-1141)	source
Qy 13 GlyLeuLeuGlyCys1leIleThrSerLeuThrGlyArgAspLysAsnGlnValGluGly 32	
Qy 33 GluvalGinIlevalSerThralaAlaGinThrPheLcuAlaThrCysIleAsnGlyVal 52	
Qy 53 CysTrpThrValTyrHisGlyalaGlyThrArgThrIleAlaSerProLysGlyProVal 72	
Oy 73 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 92 :::	
Oy 93 ArgSerleuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 112	
Oy 113 AspValileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArg 131	
Qy 132 ProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeu 145	
Oy 146 CysproalaGlyHisalaValGlyIlePhcargalaalaValCysThrArgGlyVal 164	
Oy 165 AlafysalaValAspPheIleProValGluSerLeuGluThrThrMct 180 :::	BASE COUNT 147 ORIGIN
BEG31437 LOCUS BF631437 LOCUS DEFINITION HVSMED0015P05f Hordeum vulgare seedling shoot EST 122-OCT-2001 DEFINITION HVSMED0012 (Dehydration stress) Hordeum vulgare subsp. vulgare CDNA clone HVSMED0015P05f, mRNA sequence.	Alignment Scores: Pred. No.: Score: Percent Similarity: Best Local Similarity: Ouery Match:
ACCESSION BF631437 EMERSION BF661437.2 GI:13092107 KEYMORDS EST.	DB: IIS-09-965-594-1 (1-182
W E E E	Oy 35 Gluileval Db 1 CAGCTGCC
REFERENCE 1 (bases 1 to 779) AUTHORS Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu, Henry, D., Palmer, M., Rambo, I., Simmons, J., Choi, D.W., Fenton	Qy 55 ThrvalTy1 :: Db 49 ACCGGTG
.R.D., Oates,R. and Main,D. IITLE Development of a quetically and physically anchored EST resource for harlev genemics: Morex drought-stressed seedling short char	Оу 74
library JOURNAL Unpublished	Db 109 CCCIGITCI

```
/db_xref="taxon:112509"
/db_xref="taxon:112509"
/clone="HVSMED0015p05f"
/tissue_type="Seedling shoot"
/lab_bost="Tc121" Seedling shoot
/lab_bost="Tc121" Seedling shoot
/lab_bost="Tc121" Seedling stress)"
/clone_lib="Hordeum vulgare seedling shoot EST library
/clone_lib="Hordeum vulgare seedling shoot EST library
/clone_lib="Hordeum vulgare seedling shoot EST library
/clone_loot for stress)"
/note="Vector: lambdaZAP; Site_1: ECORI; Site_2: Xhol;
/seeds were surface sterlized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedlings were
crystallization dishes. Five-day old seedlings were
incubated at 90% RH for 24 hr. Shoots were then harvested,
total RNA was prepared, poly(A) RNA was purified, one
primary unamplified cDNA library was made, 600000 pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids.
These steps were performed in the TJ Close laboratory at
the University of California, Riverside (Choi, Close,
Fenton). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
sequence for more details on library preparation and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TIGCCAGGIGGCCGCGGIGGACGCGCGCCAGGGGAAGCGGCTCGIGGICG 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpagcs/bgn/31/cover.html)* 216 c 311 g 105 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MACGGCGCCCCGCGAACAACTACAGGACGGTGC------TGCTGG 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      rHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVallle--- 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    alSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .------GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
2000 this sequence version replaced gi:11895595.
                                                                                                                                                                                                                                                                                                                                                                                                . .779
Organism-"Hordeum vulgare subsp. vulgare"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          779
47
14
63
30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                n Hall, Člemson, SC 29634, USA
556 7288
56 4293
                                        niversity Genomics Institute
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                       : AATTAACCCTCACTAAAGGG
                                                                                                                                                                                                                                                                                                                         ty sequence stop: 588.
Acation/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sequence analysis see
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2) x BF631437 (1-779)
                                                                                                                                                                                                                                                                                                                                                                                                                                                   /mol_type="mRNA"
                                                                                                                                                                                                                             ng@clemson.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21
97.50
39.87%
30.72%
10.23%
                                                                                                                                                                                                                                                               ases
```

```
Unpublished
                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                       176
                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            90
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                438
                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                            ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 15
BF307233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BF182274 846 bp mRNA linear EST 31-OCT-2000 601804028F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035102 5',
pBluescript II SK- plasmids were excised from the lambda ZAP clones by superinfection with ExAssist (Stratagenc) Abage. The library was normalized using method 4 described in Bonaldo et al. 916 Genome Research 6: 791-866.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           L Unpublished
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Lordar Hennighausen Ph.D., Robin Humphreys
CDNA Library Arrayed by: Life Technologies, Inc.
Cona Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incyte Genomics. Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9308 row: g column: 07
High quality sequence stop: 696.

Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            124
                                                                                                                                                                                                                                                                                                                 171 CACCACCATACCCTTGCTCTCACGTTGCTCACACCAAAATTATGCCCATACGGCCACA 230
                                                                                                                                                                                                                                                                                                                                                                                          284
                                                                                                                                                                                                                                                                                                                                                                                                                         93 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         405 GAAGGGGACGGGAATTCCGTTTGCGGAAAAGTGAGCGCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                          72
                                                                                                                                                                                                                                                                                                                                                         73 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         345 GAGATCATTTGCATGTGGCTTTAGTCACCCCAAGAAGAGCCCTGGGAGTGGGCATTTATAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          125 -------GlySerLeuSerProArgProlleSerTyrLeu
                                                                                                                                                                                                                                                                                       TyrHisGlyAlaGiyThrArgThrileAlaScrProLys-----GlyProVal
                                                                                                                                                                                                                                                                                                                                                                                         231 ACAAAGTTACATACACGG -----AAGGACCAGCGCGTTGGCCACCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----ProvalArgArgArgGlyAspSerArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 (bases 1 to 846)
NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          137 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 149
                                                                                                                                     701
32
14
45
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CICCAAGGCAGCAAIGGGAGCCIITCGCGGGTGIGCGGGT 503
                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                       US-09-965-594-1 (1-182) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BF182274.1 GI:11050416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          113 Aspvalile-----
                                                                                                                                   25.4
96.00
40.718
28.328
10.078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mRNA sequence.
BF182274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mus musculus
                                                                                                                                                                                     Best Local Similarity:
                                                                     ď
                                                                                                                                                                     Percent Similarity:
                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         285
                                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BF182274/C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                       ..
9
                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                   ORIGIN
                                                                                                                                   Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
BF307233 901 bp mRNA linear EST 21-NOV-2000 601891502F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4137145 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /clone="IMAGE:4035102"
/tissue_type="tumor, gross tissue"
/dev_stagg="7" months"
/lab_host="bH108"
/clone_lib="NOI_CGAP_Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Sall;
Site_2: Not!, cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Robin Humphreys,
NIH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            128 uSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeu---LeuCysPr 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            147 oAlaGlyHis---AlaValGlyIlePheArg------AlaAlaValCysThrAr 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           757 GGTTCTGCTACCAGAAACAGCCTGGATGAAGAAGGACCA-------CATCCTTC 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    709 GGTTCTTCAGTCCCAAGTGGGCTGGAGAAAGTGAACACAGAGAAAAGAGGACGGTCCCTCA 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nGlySerArg-----SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVa 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (human)
Homo sapiens
Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 901)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      59 GlyAlaGlyThrArg-ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              609 -ACTAGACACACCT -- GTAATCCAGGAGGAACGCNTGGAGGAACAGAGGGACTCC -- CT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               108 1ThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                                                                                                                                                                                                                                                     1 others
                                                                                                                                                                                                                                                                                                                                                  846
46
13
33
31
9
                                                                                                                                                                                                                                                                                                                                                                         Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                     210 t
                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-1 (1-182) x BF182274 (1-846)
                                                                                                                                                                                                                                                  241 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BF307233.1 GI:11254342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gGlyValAlaLys 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TGGGATGACTGAG 426
                                                                                                                                                                                                                                                  218 c
                                                                                                                                                                                                                                                                                                                                                  32.5
96.00
47.20%
36.80%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        mRNA sequence.
BF307233
```

us-09-965-594-1.rst

```
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be clound through the I.M.A.G.E. Consortium/LLNI. at: image.llnl.gov plate: LLCM1044 row: c column: 02
High quality sequence start: 6
High quality sequence start: 6
High quality sequence stop: 684.

Location/Cualifiers

Amol_type="mRNA" [846.]

Clone="Image: 187145"

Lissue=type="mRNA" [866."

Alb_nost-"blubb (phage resistant)"

Clone="Image: 187145"

Lissue=type="Image: 187460"

Alb_nost-"blubb (phage resistant)"

Clone="Image: 187460"

Anote-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_2: Mhol. cDNA made by oligo-di priming.

Clone="Image: 187460"

Clone="Image: 187460"

Anote-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_2: Mhol. cDNA made by oligo-di priming.

Clone="Image: 187460"

Location-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_2: Mhol. cDNA made by oligo-di priming.

Clone="Image: 187460"

Location-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_2: Mhol. cDNA made by oligo-di priming.

Clone="Image: 187460"

Location-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_2: Mhol. cDNA made by oligo-di priming.

Clone="Image: 187460"

Location-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_1: EcoR1: Site_2: Mhol. cDNA made by oligo-di priming.

Clone="Image: 187460"

Location-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_2: Mhol. cDNA made by oligo-di priming.

Clone="Image: 187460"

Location-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Site_1: EcoR1: Mhol. cDNA made by oligo-di priming.

Location-"Organ: muscle; Vector: porb7: Site_1: EcoR1: Mhol. cDNA made by oligo-di priming.

Location-"Image: 187460"

Locat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              86 pProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       106 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        rLeuLeuSerProArgProIleSerTyrLeuLysGlySerScrGlyGlyProLeuLeuCy 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66 aSerProLysGlyProValIleGlnMetTyrThrAsnValAsplysAspLeuValGlyTr 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             52 ValCysTrpThr-ValTyrHis------GlyAlaGlyThrArgThrIleAl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sProAlaGlyHisAlaValGly-----IlePheArgAlaAlaValCys 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    901
37
8
8
45
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Search completed: August 31, 2003, 04:27:18 Job time : 1773.86 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-1 (1-182) x BF307233 (1-901)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               39.5
95.50
38.46%
31.62%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      146
                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BASE COUNT
                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 17:42:58; Search time 44.1697 Seconds Run on:

(without alignments)
700.745 Million cell updates/sec

US-09-965-594-12

1 MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETIMRSP 195 Perfect score: Sequence:

Scoring table:

BLOSUM62 Gapop 10.0 , Gapext 0.5

1107863 seqs, 158726573 residues Searched:

1107863 Total number of hits satisfying chosen parameters:

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:
3: /SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:
4: /SIDSI/gcgdata/geneseqy-geneseqp-embl/AA1983.DAT:
5: /SIDSI/gcgdata/geneseqy-embl/AA1984.DAT:
6: /SIDSI/gcgdata/geneseqy-embl/AA1984.DAT:
7: /SIDSI/gcgdata/geneseqy-embl/AA1986.DAT:
8: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:
9: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:
10: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:
11: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:
12: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:
13: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:
14: /SIDSI/gcgdata/geneseqy-embl/AA1993.DAT:
15: /SIDSI/gcgdata/geneseqy-embl/AA1993.DAT:
16: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1993.DAT:
16: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1994.DAT:
17: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1994.DAT:
18: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
19: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
19: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
20: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
21: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
22: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
23: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
24: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT:
24: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1999.DAT: /SIDS1/gcgdata/geneseq/genesegp-embl/AA2003.DAT:* /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT: A_Geneseq_19Jun03:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No. SC No. 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	5core 1021 998 982 981 951 953 929	Match 100.0 97.7 96.2 96.1 93.1 93.0	Match Length DB 100.0 195 21 97.7 197 21 96.2 195 21 96.1 197 21 93.9 197 21 93.0 197 21 93.0 197 21	DB 12 21 21 21 21 21 21 21 21	AAB15220 AAB15221 AAB15221 AAB15222 AAB15222 AAB15226 AAB15224 AAB15224	Description Hepatitis C virus
68 6	899.5	88.1	191	21	AAY44728	Ú

Modified hepatitis C virus (HCV) NS3 protease comprising at least 1

WPI; 2000-465976/40. N-PSDB; AAA73329.

HCV-1 polyprotein.	HCV qenomic amino	Sequence encoded 1	Protein sequence o	Peptide encoded by	Sequence encoded i	Sequence encoded i	a)	Protein encoded by	Composite hepatiti	Amino acid sequenc	Polyprotein encode	Compiled HCV seque	Hepatitis C virus	Hepatitis C virus	HCV polyprotein.	HCV polyprotein.	Hepatitis C virus	HCV polyprotein la	Hepatitis C virus	Hepatitis C virus	A nonstructural pr	HCV-1 NS3/4a mutan	Hepatitis C virus	HCV-1 NS3/4a confo	Hepatitis C virus	HCV NS3 protein.	HCV NS4A-NS3 comp1	polypepti	HCV amino acid seq	ပ က	itis C	ပ	titis C	Hepatitis C virus	Hepatitis C virus
AAR34009	AAR40120	AAP92041	AAP90158	AAP90164	AAP92047	AAP92050	AAP90288	AAB18540	AAR70230	AAY14975	AAB18541	AAR21519	AAR31621	AAR90931	AAW34480	AAW40038	AAE22049	AAU84597	AAB15211	AAR51170	AAW31884	AAE18689	AAU76377	ABG72261	AAR66995	AAW93482	AAY24943	AAR25135	æ	AAR08123	184	066	83	AAE21838	183
14	14	10	10	10	10	10	10	21	16	50	21	13	14	17	18	13	23	23	21	15	18	23	23	24	15	20	20	13	33	1	23	23	23	23	23
2816	3011	1766	1786	2261	2301	2436	2436	2772	2894	2955	2955	3011	3011	3011	3011	3011	3011	3011	182	609	631	989	989	686	3011	631	665	2435	2436	2772	632	632	989	686	686
	87.7															•			•			•										•			
86	895.5	8	8	8	8	94	94	8	8	8	6	8	3	3	5	8	8	8	894	894	894	894	894	ů.	893.5	ω	891.5	6	6	6	890	890	850	850	890
10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	58	30	31	32	33	34	35	36	37	38	38	40	41	42	43	44	ት

ALIGNMENTS

Hepatitis, NS3 protease, viral replication; chronic liver disease, liver failure; liver cancer; mutant; mutein. Goldfarb V; Hepatitis C virus NS4A-NS3 fusion protease #2. 2hang Y, AAB15220 standard; protein; 195 AA (BRIM) BRISTOL-MYERS SQUIBB CO. 06-JAN-2000; 2000WO-US00345. Wittekind M, Weinheimer S, 99US-0115271, 19-DEC-2000 (first entry) Hepatitis C virus. Synthetic. WO200040707-A1. 08-JAN-1999; 13-JUL-2000 AAB15220; RESULT 1

```
ö
                                                                                                                                                                                                fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                 INGVCWTVYHGAGIRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLV 120
                                                                                                                                                                                                                                                                                                                                                                                                                              TRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treamments of HCV infection. This is useful as HCV can load to chook it invention concerns a number of NS3 mutants and NS3-NS4A
                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                      9
substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                          1 MKKKGSVVIVGRIVLNGAYAQQTRGEEGCQETSQTGKDKNQVEGEVQIVSTAAQTFLATC
                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protease; viral replication; chronic liver disease; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                0:
                                                                                                                                                                                                                                                                                                  Query Match 100.0%; Score 1021: DB 21; Length 195; Best Local Similarity 100.0%; Pred. No. 2.7e-98; Matches 195; Conservative 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ဗ
                                                         Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15221 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BRIM ) BRISTOL-MYERS SQUIRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first cntry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                       195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                              variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δά
                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NS3 and NS4A protease enzymas. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAR 178
                                                                                                                                                                                                                                                                                                                                                                                                                                             9
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                               1 MKKKGSVVIVGRIVLNG - - AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           59 TCINGVCWTVYHGAGTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTGGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                   ,
,
                                                                                                                                                                                                                                                                                                                                                Length 197;
                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                 Score 998; DB 21;
Pred. No. 7e-96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Repatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                 1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Υ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zhang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15212 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                               English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   179 AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97.78;
98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                 Matches 193; Conservative
                                                                               23; Fig 13; 66pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
                                                                                                                                                                                                                                                                                                                   197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAA73328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15212;
                                                                                                                                                                                                                                                                                                                       Sednence
                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                       variant.
                                                                                 Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
TRHADVIPVRRRGDSRGSI,LSPRPISYLKGSSGSPLLCPAGHAVGIFRAAVCTRGVAKAV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09
                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                   The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis. Liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
              least 1
hydrophilic
                                         U
                                                                                                                                                                                                                                                                                                                                                 INGVCWTVYHGAGIRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLV
            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophil amino acid, useful for screening inhibitors that may treat hepatitis
                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                         ć
                                                                                                                                                                                                                                                                                              Length 195;
                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                            Score 982; DB 21;
Pred. No. 3.2e-94;
0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus NS4A-NS3 fusion protease #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ą
                                                                              Example 2; Fig 10; 65pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15222 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                              / Match 96.2%;
Local Similarity 97.4%;
les 190; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                     195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W0200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15222;
                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                            21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                               Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
ò
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 අ
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                59 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MKKKGSVVIVGRINLSGDTAYAQQTRGEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MKKKGSVVIVGRIVLNG--AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                           ~
                                                                                                                                                                                                                                                                                                                                                                                         Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                           4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                         Score 981; DB 21;
Pred. No. 4.1e-94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                                         1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                  Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15226 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   179 AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                           96.1%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Weinheimer
                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity
Matches 190; Conserv
                                                                                                                                                                                                                                                                                                                                                          197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:Inhibitor complexes. This sequence contains the alpha-helix0 wild-type sequence. Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C Example 5; Fig 18; 66pp; English. 197 AA; Sequence

118 61 TCINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120 121 LVTRHADVIPVRRRGDSRGSILSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180 28 9 1 MKKKGSVVIVGRIVLNG--AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFILA TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDI,VGWPAPQGSRSLTPCTCGSSDLY LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK Gaps 2: Length 197; Indels Score 959; DB 21; Pred, No. 8.2e-92; 1; Mismatches 6; AVDFIPVESLETTMRSP 195 AVDFIPVESLETIMRSP 197 93.98; 95.48; Best Local Similarity 95.4 Matches 188; Conservative 59 119 179 181 Query Match a ò g ò g ò g

AAB15223 standard: protein; 197 19-DEC-2000 (first entry) AAB15223;

Hepatitis C virus NS4A-NS3 fusion protease #5.

Hepatitis; NS3 protease; viral replication; chronic liver disease; 11ver failure; liver cancer; mutant; mutein. Hepatitis; NS3

Hepatitis C virus Synthetic.

13-JUL-2000

WO200040707-A1

2000WO-US00345 06-JAN-2000; 08-JAN-1999;

(BRIM) BRISTOL-MYERS SQUIBB

Goldfarb V; Zhang Y, Weinheimer S, Wittekind M,

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-liveriant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDIAYAQQTRGEEGCQETSQIGRDKNQVEGEVQIVSTAIQIFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRIVLNG--AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 951; DB 21;
Pred. No. 5.6e-91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred, No. 5.6e
1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ę
                                                                                                                                                       23; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15224 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       179 AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best_Local Similarity 94.9
Matches 187; Conservative
2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                           197 AA;
                    N-PSDB; AAA73332.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15224;
                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                         Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULI 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB1522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

118

28

Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein. Hepatitis C virus NS4A-NS3 fusion protease #6. Hepatitis C virus WO200040707-A1 13-JUL-2000 Synthetic THE STATE OF STATE OF

06-JAN-2000; 2000WO-US00345 08-JAN-1999; Goldfarb V; Zhang Y, Weinheimer S, Wittekind M,

(BRIM) BRISTOL-MYERS SQUIBB CO.

```
The present sequence is a mutated version of a fusion protein created using the Hepatitls C virus (HeV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and proteins the alpha-helixor?
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TCINGVCWTVYHGAGTRTIASPKGPV1QMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LVTRHADVIPVRRRGDSRGSLI,SPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis \mathbb C
                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MKKKGSVVIVGRIVLNG--AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                    Score 939; DB 21; Length 197;
Pred. No. 1e-89;
1; Mismatches 8; Indels 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                   3,
                                                                                                                                  Claim 23; Fig 16; 56pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15225 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                         92.0%;
93.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                     184; Conservative
               2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                            197 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                             N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15225;
                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                variant
                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as NCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NS3 catalytic domain; NS4A peptide; NS4A-NS3 fusion construct; diagnosis; serine protease; trypsin family; screening; anti-viral compound; treatment; inhibitor; therapeutic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 MKKKGSVVIVGRIVLNG--AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1..12
/label= NS4A_peptide_1
/label= "Covalently attached to amino terminus of NS3
catalytic domain"
                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "Wild type Proline is replaced with Lysine"
                                                                                                                                                                                                                                                                                                                                                                                                                                  5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 catalytic domain fusion protein-1.
                                                                                                                                                                                                                                                                                                                                                                                                 Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                 Score 929; DB 21;
Pred. No. 1.1c-88;
                                                                                                                                                                                                                                                                                                                                                                                                                                 3; Mismatches
Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ź
                                                                                                                                              English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY44728 standard; Protein; 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                91.0%;
92.9%;
ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-MAY-2000 (first entry)
Weinheimer
                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 183; Conservative
                                                                                                                                             23; Fig 17; 66pp;
                            WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                  Rest Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                       197 AA;
                                          N-PSDB; AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200001718-A2
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY44728;
                                                                                                                                                                                                                                                                                                                                                                          Sednence
                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                            variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Peptide
                                                                                                                                              Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY44728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                rυ
συ
                                                                                                                                                                                                                                                                                                                                                                                                                  2; Gaps
```

```
112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                 This sequence represents the entire hepatitis C virus polyprotein. HCV is a member of the flavivirus family and appears to encode a basic polypeptide domain ("C") at the N-terminal of the viral polyprotein, followed by two glycoprotein domains ("EI", "EZ/NSI"), upstream of the nonstructural genes NS2 through NS5. See also AAQ39134-48, AAR33982-4008 and AR38088-89.
(Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     52 AAQIFLATCINGVCWTVYHGAGIRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1125 GGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV genomic amino acid sequence isolated from infected human LG.
                                                                                                                                                                                                                        Immuno-reactive hepatitis C virus polypeptide compsns. - contg. at least 2 sequences from the first variable domain of distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 14; Length 2816;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV; NANBHV; human growth hormone; HGH; secretion signal; fusion protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88.0%; Score 898.5; DB 14; larity 85.8%; Pred. No. 6.1e-84; Conservative 9; Mismatches 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      $
                                                                                                                                                                                                                                                                                               Disclosure; Fig 9; 106pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR40120 standard; Protein; 3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93WO-0S00907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92US-0830024
                                 92WO-US07683
                                                                       91US-0759575
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (updated)
(first entry)
                                                                                                                                                 Weiner AJ;
                                                                                                                                                                                 WPI; 1993-117468/14.
                                                                                                            (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 175; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2816 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29-JAN-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         W09315193-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31-JAN-1992;
                                                                                                                                                                                                                                                              HCV isolates
                                   11-SEP-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25-MAR-2003
                                                                                                                                                 Houghton M,
01-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAR40120;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAR40120
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        \ddot{o}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                   The present protein sequence is the fusion polypeptide, comprising the hepatitis C virus NS4A poptide-1 fragment, covalently attached to the amino terminus of NS3 catalytic domain. This fusion polypeptide to the NS3 domain expressed in a stable, soluble form. This facilitates the use of the polypeptide in direct screening of potential anti-viral compounds, that are used for diagnosis and treatment of hepatitis C virus infection. It is also used to screen for inhibitors of serine protease activity. The polynucleotides are also useful to identify diagnostic or therapeutic compounds and for recombinant production of the fusion
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHAD 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      126 VIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF1PV 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SVVIVGRIVLNGAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1; Gaps
                                                                                                                                                                                                                                        Novel polypeptide comprising hepatitis C virus NS4A and NS3 domains, useful for screening for compounds useful for the diagnosis and treatment of hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus; PCV; asymptomatic; chronically infected; epitope; vital isolate; domain; immunological; cross reactive; envelope protein; vaccine; qp53(BVDV)/qp55; hog cholera virus; pestivirus; NS1; flavivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 899.5; DB 21; Length 191;
Pred. No. 1.3e-85;
6; Mismatches 8; Indels 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAR34009 standard; Protein; 2816 AA
                                                                                                                                                                                                                                                                                                                   Claim 4; Fig 2; 30pp; English
                                                                                          98US-0091675
                                                     99WO-US15035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 92.13
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ENLETTMRSP 190
                                                                                                                                                                   Bukhtiyarova
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (updated)
                                                                                                                              (UYFL ) UNIV FLORIDA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV-1 polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            191 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9306126-A1
                                                     02-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-2003
26-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polypeptide
               13-JAN-2000
                                                                                                                                                                 Dunn BM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAR34009;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
```

q

õ

q ö

ò g ò

1064

Indels

Sun Aug

```
Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP90158 standard; protein; 1786
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87US-0122714.
87US-0139886.
88US-0161072.
                                                                           Kuo G;
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus; vaccine
                                                                                                                                                                                                                                                                                                          Best Local Similarity 85.8
Matches 175; Conservative
                                                                          Houghton M, Choo QL,
                                                                                              WPI; 1989-159274/22.
                                                                                                                                                                                                                                                                             1766 AA;
                                                                                                          N-PSDB; AAN92097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               an troglodytes.
                                                     (CHIR ) CHIRON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-MAR-2003
10-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26-JUL-1989
                                                                                                                                                                                                                                                                             Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          469
                                                                                                                                                                                                                                                                                                                                                                                  23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAP90158;
                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAP90158
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                      .;
?
                                                                                                                                                                                                                                                                                                                                                                                                                Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 141, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                            RNA was isolated from the plasma of a HCV scropositive human (designated "LG") and cDNA was prepared from it. The cDNA was PCR amplified using specific primers with sequences based on the prototype HCV-1 cDNA sequence (GENBANK M62321). Further amplification using nested primers resulted in 7 adjacent HCV DNA fragments which could be assembled into a full-length sequence. The DNA sequence was determined and translated into the genomic amino acid sequence. Comparison of the LG genomic amino acid sequence with that from HCV-1 showed 134 amino acid differences. (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                   87.7%: Score 895.5; DB 14; Length 3011; 85.8%; Pred. No. 1.4e-83; Live 8; Mismatches 10; Indels 11;
                                                                                    id pHCV-162 is a mammalian expression systems for HCVI
- useful for diagnosing HCV infection and as vaccines
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                Frail DE;
                                Devare SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 1; Page 39-49; 100pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAP92041 standard; protein; 1766 AA
                                Desai SM,
                                                                                                       for preventing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87US-0122714.
87US-0139886.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003 (updated)
02-MAR-1990 (first entry)
                                                                                                                                                                                                                                                                                                Best Local Similarity 85.8 Matches 175, Conservative
                                          Zeck BJ;
                                                                                    plasmid pHCV-162
                                                              WPI; 1993-258673/32.
                                Casey JM,
                                                                                                                                                                                                                                                                 3011 AA;
           (ABBO ) ABBOTT LAB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-NOV-1987;
30-DEC-1987;
                                          Yamaguchi J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP92041;
                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                      Query Match
                                SI,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11

MAPP 2041

IND AAPP 2041

XXX

XXX

DDT 25-1

DDT 25-1

DDT 25-1

AXX

Hep Seq 141

EP 31-1

PF 18-1

PR XXX

PF 18-1

PR 30-1
                                                                                                                                                                                                                                                                                                                                                 g
 ò
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                \tilde{\varsigma}
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
```

```
It is the sequence encoded in the open reading frame of hepatitis C virus cDNA inserts in clones 141,m 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, a33f, 33g and 39c. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   289 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVBGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVG---RIVLNG-----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                              Claim 13; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 10; Length 1766;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Protein sequence of hepatitis c virus composite cDNA
Purified hepatitis C virus
- and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87.6%; Score 894.5; DB 10;
85.8%; Pred. No. 8.6e-84;
tive 8; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 172 CTRGVAKAVDFIPVESLETTMRSP 195
```

œ

7

11;

Indels

Length 2261;

XXX XXX BET X BET

```
180 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLIGCIITSLTGRDKNQVEGEVQIVST 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 12f through 15e.
                                                                                                                                                                                                                                                                                                                                                                                                                                 AAQTFLATCINGVCWTVYHGAGIRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                      The sequence is the peptide encoded by the composite hepatitis C virus (HCV) CDNA of AAN90331. The polypeptides are used to diagnose HCV-induced NANH, to raise antibodies for immunoassay or treatment, cur produce vaccines.

(Updated on 25-MAR-2003 to correct PR field.)
                                                          Hepatitis C virus gene - used for prodn. of polynucleotide polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                      DB 10;
                                                                                                                                                                                                                                                                                  Ouery Match 87.6%; Score 894.5; DB 10; Best Local Similarity 85.8%; Pred. No. 1.2e-83; Matches 175; Conservative 8; Mismatches 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            172 CTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAP92047 standard; protein; 2301 AA.
                                                                                                                            Disclosure; fig 32; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (CHIR ) CHIRON CORP
               WPI; 1989-215054/30
                                                                                                                                                                                                                                                      2261 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                               N-PSDB; AAN90331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP92047;
                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
G
                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
 ô
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                               3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus gene · used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                   The sequence is encoded by the composite cDNA of AAN90327. These antigens react with antibodies in patients with non-A non-B heptilis (NANBH). They can be used to diagnose HVV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVG---RIVLNG-----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                               11; Gaps
                                                                                                                                                                                                                                                                                                                                                              Query Match 87.6%; Score 894.5; DB 10; Length 1786; Best Local Similarity 85.8%; Pred. No. 8.7e-84; Matches 175; Conservative 8; Mismatches 10; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus; clone 12f; clone 15e; probe: vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Peptide encoded by composite hepatitis C virus CDNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAP90164 standard; protein; 2261 AA
                                                                                                                                                                                                         Disclosure; fig 26; 30pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                              ::
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (CHIR ) CHIRON CORPORATION.
88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Choo QL, Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (updated)
(first entry)
                               (CHIR ) CHIRON CORPORATION.
                                                              Choo OL,
                                                                                          WPI; 1989-215054/30.
N-PSDB; AAN90327.
                                                                                                                                                                                                                                                                                                                                   Sequence 1786 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-NOV-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAR-2003
01-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26-JUL-1989
                                                            Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAP90164;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    469
```

QQ

g Š a qq

õ

N-PSDB; AAN92103.

```
It is the sequence encoded in the open reading frame of hepatitis C virus (HCV) cDNM inserts in clones 12f through 15e. It is antiqenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.

(Updated on 25-WAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVG---RIVING-----AYAQQTRGEEGCQETSOTGRDKNOVEGEVQTVST 51
                                                                                                                                                                                                                                                   Query Match 87.6%: Score 894.5; DB 10; Length 2301; Best Local Similarity 85.8%; Pred. No. 1.2e-83; Matches 175; Conservative 8; Mismalches 10; Indels 11; Gaps
                              Purified hepatitis C virus
- and associated nucleic acids and polypeptide(s)
                                                                           Claim 13; Figure 32-1 - 32-7; 139 pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                        Sequence 2301 AA;
q
                                                                                                                                                                                                                                                                                                                                                                                             QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                        ολ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

Search completed: August 30, 2003, :9:12:22 Job time: 47.1697 sccs

CIRCVAKAVDFIPVENLETIMRSP 583

```
GenCore version 5.1.6 Copyright (c) 1993 \cdot 2003 Compugen Ltd.
```

OM protein - protein search, using sw model Run on:

August 30, 2003, 19:02:22; Search time 16.0488 Seconds (without alignments) 1168.492 Million cell updates/sec

US-09-965-594-12 1021 1 MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETIMRSP 195

Title: Perfect score:

Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308

283308 segs, 96168682 residues

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Description	qenome polyprotein							genome polyprotein	genome polyprotein			polyprotein - marm	polyprotein - dour	<u>п</u>	_	angio-associated m	nitrate/nitrite se	nitrate/nitrite se	env polyprotein -	insulin-like growt	hypothetical prote	sur	probable aromatic	beta transducin ho	probable periplasm	centromere protein	hypothetical prote	_	·H	
	1.0	GNWVC3	S40770	GNWVCH	GNWVTW	CNWVTC	GNWVCJ	A45573	S18030	JC5620	JQ1303	GNWVJB	T08839	T08841	B71360	AB1775	139383	B81104	C81911	VCMVFG	IOHOI	T49806	A46165	H83144	T18234	B71284	S28261	T34513	П	E95261	
	DH	Н	-	-	-	٦	_		П	-							7	7	7		-	~	7	7	7					7	
	Length	3011	3011	3011	3010	3010	3010	3010	3010	3014	3033	3033	2970	3005	495	1334	452	290	290	642	259	354	404	209	981	398	2663	3507	270	393	
	Match	7	~	86.1	\sim	82.5	82.1	ď.		74.5		Š.	25.5	25.3	8 . ₹		8.0			7.8	•	٠	•	7.7	•	٠	٠	7.6	7.5	7.5	
	Score	894.5	893.5	879.5	847	842	838	837	821	191	670.5	668.5	260	258	85.5	ď.	82		80.5	79.5	79	79	~			77.5			76.5	76.5	
+[00	Nesure No.	7	7	æ	4	S	9	7	6 0	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53	

serine proteinase	hypothetical prote	probable beta-qluc	hypothetical prote	env polyprotein pr	orsellinic acid sv	zona pellucida qly	probable exported	hypothetical prote	ice nucleation pro	serine proteinase	conserved hypothet	env polyprotein	hypothetical prote	phosphate ABC tran
B98127	C/2/03 A84782	T35785	S41034	VCMVSA	T30871	S70401	AH1030	523441	T18042	99668н	G87265	VCFVER	S76618	Н84203
C) C	7 ~	7	a	-	7	~	7	7	7	7	7	~	~	7
397	574	859	882	639	1293	415	492	755	1176	239	239	603	317	266
7.5	7 . 4	7.4	7.4	7.3	7.3	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2
76.5	76	75.5	75.5	7.5	75	74.5	74.5	74.5	74.5	74	74	74	73.5	73.5
30	32	33	34	35	36	37	38	33	40	41	42	43	44	45

ALIGNMENTS

RESULT 1 GENWUZ3 GENWUZ3 GENWUZ3 GENWUZ1 M.COUTALEL 30-Sep-1392 requested in Cartal protein N44b, nonstructural protein N55 C.Species hepatitis C virus C.Species hepatitis C virus C.Species hepatitis C virus G.Species hepatitis C virus G.Species hepatitis C virus G.Species hepatitis C virus R.CDOG, O.L.; Richman, K.H.; Han, J.H.; Berger, K.: Lee, C.; Dong, C.; Gallegos, C.; R.CDOG, O.L.; Richman, K.H.; Han, J.H.; Berger, K.: Lee, C.; Dong, C.; Gallegos, C.; R.CDOG, O.L.; Richman, K.H.; Han, J.H.; Berger, K.: Lee, C.; Dong, C.; Gallegos, C.; R.CDOG, O.L.; Richman, K.H.; Han, J.H.; Berger, K.: Lee, C.; Dong, C.; Gallegos, C.; R.CDOG, O.L.; Richman, K.H.; Han, J.H.; Berger, K.: Lee, C.; Dong, C.; Gallegos, C.; R.COGS contain Cold Cold A.R.CESS cold Cold B.R.CESS cold Cold A.R.CESS cold A.R.CESS cold A.R.CESS cold
--

us-09-965-594-12.rpr

```
C; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A:Cross-references: GB:D00831; NID:g221511: PIDN:BAA00705.1; PID:g221512
A:Experimental source: isolate HC-J1
A:Experimental source: isolate HC-J1
A:Experimental source: isolate HC-J1
C:Superfamilty: hepatitis C virus genome polyprotein
C:Superfamilty: hepatitis C virus prodicted <CPC>
F:116-191/Product: capsid protein M #status predicted <MEE>
F:116-191/Product: morstructural protein NSI #status predicted <NSI>
F:330-123/Product: nonstructural protein NSI #status predicted <NSI>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif A (P-loop)
F:1316-1862/Product: nonstructural protein NS4 #status predicted <N4A>
F:1861-1862/Product: nonstructural protein NS5 #status predicted <N4B>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NSS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A;COSS-references: EMBL:D10749; NID:9221586; PIDN:BAA01582.1; PID:9221587
A;COSS-references: EMBL:D10749; NID:9221586; PIDN:BAA01582.1; PID:9221587
B;OXamoto, H.; OXada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
A;Recoule type: genomic RNA
A;Rolecule type: genomic RNA
A;Residues: 1-513 <OX2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein – hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1065 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIOMYTNVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                                                                    .005 RRGREILLGPAIXGWSKGWRLLAPITAYAQQTRGLLGCI1TSLTGRDKNQVEGHVQIVST 1054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    112 CGSSDLYLVTRHADVIPVRRRCDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                         52 AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVOKBLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                         112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C.Species: hepatitis C virus
C.Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C.Accession: S40770; PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCOETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 87.5%; Score 893.5; DB 1;
85.8%; Pred. No. 2.1e-74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RickAmoto, H.
submitted to the EMBL Data Library, March 1992
A;Reference number: $40770
A;Accession: $40770
A;Accession: $40770
A;Accelle type: genomic RNA
A;Residues: 1-3011 <0KA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1185 CTRGVAKAVDFIPVESLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                       CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                         CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity 85.8 75; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 175;
                                                                                                                                                                                                                                                                                         172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                      q
                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                      g
                                                      ò
                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                      ò
```

```
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4 *status predicted <NNB>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NNS>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NNS>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240
qenome polyprotein - hepatitis C virus (strain H)
N/Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4a; nonstructural protein NS5
                                                                                                                                                                                                                               A) Description: Genomic structure of the human prototype strain H of hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                      C.Species: hepatitis C virus
A:Note: host Homo sapiens (man)
C.bate: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C.Accession: A56814; A41546
R:Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52 AAQTELATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVG---RIVLNG-----AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R;Chen, P.J.; Lin, M.H.; Tal, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.1%; Score 879.5; DB 1
83.8%; Pred. No. 4.1e-73;
live 9; Mismatches 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CIRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                         submitted to GenBank, July 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 83.8 nes 171; Conservative
                                                                                                                                                                                                                                                             A;Reference number: A36814
A;Accession: A36814
```

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δý
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                             a
                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Title: Structure and organization of the hepatitis C virus genome isolated from human A;Reference number: A38465, MUID:91140698; PMID:1847440
A;Recession: A38465
A;Accession: A38465
A;Coss-references: EMBL:368335; NID:9329770; PIDN:AAA72945.1; PID:9329771
C;Superfamily: hepatitis C virus genome polyprotein: 91ycoprotein; hydrolase; nonstructural C;Keywords: ATP; capaid protein; envelope protein: 91ycoprotein; hydrolase; nonstructural protein B;Status predicted <APA;
F;16-191/Product: envelope protein B;Status predicted <APA;
F;390-729/Product: nonstructural protein NS1 *status predicted <NS2>
F;300-1006/Product: hepacivirin *status predicted <NS2>
F;1230-1237/Region: nucleotide-binding motif B
F;1312-1317/Region: nucleotide-binding motif B
F;1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1616-1862/Próduct: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2034,250,205,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructut protein NS4) by NS5 protein NS5 protein NS5 c. Species: hepatitis C virus C virus C; Species: hepatitis C virus C; Date: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001 C; Date: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001 C; Takamizawa, A ; Moni, C ; Fuke, I ; Manabe, S ; Murakami, S ; Fujita, J ; Onishi, E ; J VIrol. 65, 1105-1113, 1991
                                                                                                                                                                      ACCOSS-references: GB.M84754
CS.Superfamily: hepatitis C virus genome polyprotein
CS.Superfamily: hepatitis C virus genome protein
CS.Superfamily: hydrolase; nonstructura
FS.116.191/Product: envelope protein
FS.116.191/Product: major envelope protein
FS.390.729/Product: nonstructural protein
FS.300.1006/Froduct: nonstructural protein NS1 *status predicted <NS2>
FS.1230-1237/Region: nucleotide-binding motif A (P-loop)
FS.1317/Region: nucleotide-binding motif B
FS.1316.1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1863-2013/Product: nonstructural protein NS4b *status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A,Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRC 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .;
O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 842; DB 1;
Pred. No. 1.3e-69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83.0%; Score 847; DB 1; 88.2%; Pred. No. 4.3e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82,5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 88.2%
Matches 157; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;1316-1319/Region: DEXH motif
                                                                                                            A; Molecule type: genomic RNA
A; Residues: 1-3010 <CHE>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                            A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

```
A.Experimental source: Japanese isolate
C.Comment: The cleavage sites of this polyprotein have not been determined.
C.Comment: The cleavage sites of this polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein;
C.Superfamily: hepatitis C virus genome polyprotein;
C.Superfamily: hepatitis C virus genome polyprotein;
F.2-115/Product: capsid protein C *status predicted <CPC>
F.116-191/Product: envelope protein M *status predicted <MEE>
F.130-139/Product: nonstructural protein NSI *status predicted <NSI>F.730-1006/Product: nonstructural protein NSI *status predicted <NSI>F.730-1006/Product: hepativirin *status predicted <NSI>F.1310-1317/Region: nucleotide-binding motif A (P-loop)
F.1315-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R.Rato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Sh
Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A,Title: Molecular cloning of the human hepatitis C virus genome from Japanese patie
A,Reference number: A39253; MUID:91088550; PMID:2175903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               nonstruc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1865-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA
A; Residues: 1-3010 <a href="https://doi.org/10.10/10/201610">https://doi.org/10.10/201610</a>
A; Cross-references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
R; Kato, N.; Ohkoshi, S.; Shimotchno, K.
Proc. Jpn. Acad. 65b, 219-223, 1989
A; Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence A; Reference number: PS0085
A; Accession: PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
   ò
                                                                                                                                                                                                         78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                       11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                              138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E;
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 30-Jun-1992 *sequence_revision 30-Jun-1992 *text_change 19-Jan-2001
C;Accession: A33253; PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                       18 AYAQQIRGEEGCQEISQIGRDKNQVEGEVQIVSIAAQIFLAICINGVCWIVYHGAGIRII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
   Gaps
   ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82.1%; Score 818; DB 1; Length 3010;
85.4%; Pred. No. 3e-69;
Live 16; Mismatches 10; Indels
   10; Indels
   Mismatches
14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA A; Residues: 2650-2707 <KA2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    al Similarity
152; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loca
Matches
```

```
Cichamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A.Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A.Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A.Accession. JC5620
A.Molecule type: mRNA
A.Molecule type: mRNA
A.Molecule type: mRNA
A.Coss references: GB:X13184
A.Escidues: 1-3014 <CHA>
A.Coss references: GB:X13184
A.Escidues: 1-3014 <CHA>
A.Coss references: GB:X13184
A.Escidues: 1-5014 <CHA>
A.Coss references: GB:X13184
A.Coss references: GB:X13184
A.Notos: the translation of the nucleotide binding: P-loop; polyprotein; selt:2.Inforduct: capsid protein A **status predicted <CPC>
F.116-191/Product: envelope protein M **status predicted <EPM>
F.116-191/Product: major envelope protein E **status predicted <AEPM>
F.116-191/Product: major envelope protein E **status predicted <AEEPM>
F.116-191/Approduct: major envelope protein B **status predicted <AEPM>
F.116-191/Approduct <AEPM **AEMM *
                                                                                                                                                                                 A.Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 3 as Trp, and TTC for residue 771 as Scr
A.Note: sequence extracted from NCB1 backbone (NCBIN:121747, NCBIP:121748)
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus grandicted common polyprotein
C.Superfamily: hepatitis C virus grandicted common polyprotein
C.Superfamily: hepatitis C virus grandicted common polyprotein; se F:2-115/product: capsid protein M *status predicted common polyprotein; se F:116-191/Product: major cnvelope protein E *status predicted common postructural protein NS2 *status predicted common postructural protein NS2 *status predicted common voleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B (P-loop)
F:1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructorian NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus
C; pate: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C; Accession: JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1616-1862/Product: nonstructural protein NG4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NG4b *status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                    'V',623-624,'S',626-652,'DL',655-761,'T',763-782 <HOW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F:390-730/Product: nonstructural protein NSI #status predicted <NSI>F:731-1007/Product: nonstructural protein NSI #status predicted <NS2>F:731-1007/Product: hepacivirin #status predicted <NS3>F:1008-1616/Product: hepacivirin #status predicted <NS3>F:1211-1218/Region: nucleotide-binding motif A (P-loop)
F:1313-1318/Region: nucleotide-binding motif B
F:1317-1320/Region: DEXH motif
                                                                                                                                            A; Note: this sequence is inconsistent with the nucleotide translation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 821; DB 1;
Pred. No. 1.1e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80.4%;
85.4%;
A; Molecule type: genomic RNA
A; Residues: 1-547,'T',549-621,'\
A; Cross-references: EMBL:X61591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 85.4%
Matches 152; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  \ddot{o}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A:Status: prefiminary
A:Status: prefiminary
A:Molecule type: DNA
A:Cross references: GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g221613
A:Cross references: HCV-JT
A:Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
F:115-Product: capsid protein C *status predicted <RED>
F:116-191/Product: capsid protein M *status predicted <RED>
F:182-193/Product: major envelope protein B: *status predicted <NSI>
F:300-1006/Product: nonstructural protein NSI *status predicted <NSI>
F:1007-1615/Product: nonstructural protein NSI *status predicted <NSI>
F:1007-1615/Product: nonstructural protein RSI *status predicted <NSI>
F:1112-1117/Region: nucleotide-binding motif A (P-loop)
F:1117/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A Variety: isolate JKI
C;Date: 19-May-2000 #sequence_revision 15-May-2000 #text_change 23-Mar-2001
C;Accession: S18030; 333570; A48332; S18029
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
Submitted to the EMBL Data Library, September 1991
A;Description: A Whole genome of hepatitis C virus cDNA was isolated from a single patie
A;Accession: S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (isolate JK1)
N'Contains: appsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
                          genome polyprotein - hepatitis C virus (strain JT)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructurotrotein NS44; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May:2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A.Cross-references: EMBL:X61596; NID:q59478; PIDN:CAA43793.1; PID:q59479
A.Experimental source: isolate JK1 from an individual
R:Honda, M.; Raneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A.Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A.Reference number: A48332; MUID:93119270; PMID:8380322
                                                                                                                                                                                                                                                                                                                                        R:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,
Virus Res. 23, 39-53, 1992
A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier:
A:Reference number: A45573; MUID:92295714; PMID:1318627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         78 ASPKGPVIQMYTNVDKDI,VGWPAPQGSRSI,TPCTCGSSDI,YLVTRHADVIPVRRKGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18 AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSTAAQIFLAICINGVCWTVYHGAGIRII 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1151 SLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Match 82.0%; Score 837; DB 1; Local Similarity 86.5%; Pred. No. 3.7e-69; es 154; Conservative 13; Mismatches 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Molecule type: genomic RNA A; Residues: 1-3010 <HON>
                                                                                                                                                                                                                                                                                    C; Accession: A45573
R; Tanaka, T.; Kato.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
procession: Adoles inconstructural protein No. 1 inconstructural protein No. 2 (5) Species: hepatitis C virus
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: Ad0250; PQ0397; PQ0559
B; Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, Virology 188, 331-341, 1992
A; Title: Full-length sequence of a hepatitis C virus genome having poor homology to A; Reference number: Ad0250; MUID:92230232; PMID:1314459
A; Accession: Ad0250
A; Molecule type: genomic RNA
A; Residues: 13:033 Accession: Ad0250
A; Molecule type: Genomic RNA
A; Residues: 13:033 Accession: Billo988; GB:D01221; NID:9221608; PIDN:BAA01761.1; PID:9221609
B; Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 131-1141, 1992
A; Reference number: PQ0393; MUID:92268871; PMID:1316939
A; Accession: P00397
                                                                          genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A Molecule type: genomic RNA
A Residues: 2578-278-74 CKHA
A Cross-references: DDBJ:D10134
B Cochem. Biophys. Res. Commun. 181, 279-285, 1991
B Cochem. Biophys. Res. Commun. 181, 279-285, 1991
A Crocssion. PQ0554; MUID:92068204; PMID:1720309
A MACCESSION: RNA
A MAINSA CROSS COMMANA
A Residues: 2678-2729 CKAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F;1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 *status predicted <NS5>
F;2018-303,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,20:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1010 LGREVLLGPADGYISKGWKLLAPITAYTQQTRGLLGAIVVSLTGRDKNEQAGQVQVLSSV 1069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10 VGRIVLNG-------AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65.5%; Score 668.5; DB 1;
llarity 62.6%; Pred. No. 1.8e-53;
Conservative 25; Mismatches 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Æ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.190 ARGVAKSIDFIPVESLDVATRTP 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein - marmoset hepatitis GB virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      173 TRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1320-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 127; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Mcetestion: J01303
A; Molecule type: genomic RNA
A; Residues: 1-3013 < CNA
A; Cross-references: GB:D00944; NID:g221650; PIDN:BAA00792.1; PID:g221651
A; Experimental source: isolate HC-J6 from a Japanese individual
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinsec; ransm
C; Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinsec; F; 2-115/Product: envelope protein M *status predicted < CNE>
F; 116-191/Product: nonstructural protein NS1 *status predicted < CNE>
F; 390-733/Product: nonstructural protein NS2 *status predicted < CNE>
F; 1310-1323/Region: nucleotide-binding motif B
F; 1320-1323/Region: nucleotide-binding motif B
F; 1320-1323/Redion: nonstructural protein NS4 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *status predicted < CNA>
F; 1867-2017/Product: nonstructural protein NS5 *SSB, S78, 627, 649, 1091, 1217, 1259, 2038, 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (isolate HC-J6)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructut protein NS4s; nonstructural protein NS5secies: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 19-May-2000 *sequencc_revision 19-May-2000 *text_change 17-Nov-2030
C;Accession: JG1303 *Sequencc_revision 19-May-2000 *text_change 17-Nov-2030
C;Accession: JG1303 *Sequence_revision 19-May-2000 *text_change 17-Nov-2030
C;Accession: JG1303 *Sequence of the genomic RNA of hepatitis C virus isolated from a hum A;Reference number: JG1303; MUID:92044440; PMID:1658196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1032 AYAQQTRGVLGAIVI,SI,TGRDKNEAEGEVQFI,STATQTFI,CICINGVMTLFHGAGSKTL 1091
                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                1092 AGPKGPUVQMYTNVDKDLVGWPSPPGKGSLIRCICGSADLYLVTRHADVIPARRGDTRA 1151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 113 GSSDLYI,VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVC 172
                                                                                                                                                                                                                                                                                                                                                                                              78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA 52
                                                                                                                                                                                                                                                     18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSIAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              152 SLLSPRPISYLKGSSGGPIMCPSGHVVGVFRAAVCTRGVAKALEFVPVENLETTMRSP 1209
                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;2015-3014/Product: nonstructural protein NS5 *status predicted <NS5> F;2210-2249/Region: interferon sensitivity determining *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17;
                                                                                                                                                                                      .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                        Query Match 74.5%; Score 761; DB 1; L
Best Local Similarity 77.5%; Pred. No. 4.3c-62;
Matches 138; Conservative 19; Mismatches 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Pred. No. 1.1e-53; 27; Mismatches 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65.7%; Score 670.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1190 SRGVAKSIDFIPVETLDIVTRSP 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10 VGRIVLNG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 126; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
```

Sun Aug 31 15:59:06 2003

ဟ 🗷

Length 495;

```
A;Residues: 1-495 <CCL>
A;Cross-references: GB:AE001199; GB:AE000520; NID:g3322402; PIDN:AAC65137.1; PID:g332
A;Experimental source: strain Nichols
                                                                                                                                                                                                                                            RiFraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T they, L.; Weidman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete. A;Reference number: A71250; MUID:98332770; PMID:9665876
                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Accession: B71360
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
                                                                                                                                 hypothetical protein TP0136 - syphilis spirochete
C.Species: Trepomena pallidum subsp. pallidum (syphills spirochete)
C.Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C.Accession: B71360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match

8.4%; Score 85.5; DB 2;
Best Local Similarity 23.6%; Pred, No. 3.6;
Matches 52; Conservative 19; Mismatches 78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: DNA
A; Residues: 1-1334 <GLA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Gene: lin2744
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Gene: TP0136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ΩΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
C;Species: marmoset hepatitis GB virus A
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1969 #text_change 17-Nov-2006
C;Accession: 108839
R:Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Mostes, C.C.; Mushahwar, I.K.
J. Gen. virol. 79, 41-45, 1998
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkcys.
A;Title: Genomic analysis of fwo GB virus A variants isolated from captive monkcys.
A;Title: Genomic analysis of MUD:98120818; PMID:9460920
A;Title: Genomic analysis of MUD:98120818; PMID:9460920
A;Stellus: translated from GD/EMBL/DDBJ
A;Residues: 1-2970 <ERK>
A;Residues: 1-2970 <ERK>
A;Cossion: Townslated from GD/EMBL/DDBJ
A;Cossion: Townslated from GD/EMBL/DDBJ
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C.C.; Mushahwar, I.K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein - douroucouli hepatitis GB virus A C;Species: douroucouli hepatitis GB virus A C;Step-1999 #text_change 17-Nov-2000 C;Accession: T08841 A:Pi D98 A:Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A:Teference number: 216486; MUID:99120818; PMID:9460920 A:Recession: T08841 A:Reference T08841 A:Status: translated from GB/EMBL/DDBJ A:Molecule type: mRNA A:Residues: 1-3005 CEBK> A:Residues: 1-3005 CEBK> A:Residues: 1-3005 CEBK> A:Cross-references: EMBL:AF023425; NID:9288599; PIDN:AAC40502.1; PID:92828600; C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1075 KCGTQSVWCIRN--DGALCHGRLSKLVELDLPTEISDFRGSSGSPILCDEGHVVGMM-VS 1131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                955 RRCDEVLVGVFPSVRALPPGFVPTAPVVVMQRGLGFFSVVKTSMLGRDERFHEGSIVVI,G 1014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCKCQPQGVWVI------RND--GALCHGTLGRTVELOLPARLCDFRGSSGSPIL 1110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            51 IAAQIFLATCINGVCWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPC 110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             111 TCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPULCPAGHAVGIFRAA 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VSTAAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDIJVGWPAPQGSRSLT 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       109 PCTCGSSDLYLVTRHADVIPVRRGDSRGSLLS------PRPISYLKGSSGGPLL 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVG-----RIVLNG-----AYAQQTRGEEGCQETSQIGRDKNQVEGEVQIVS 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3005;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 2970;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CPAGHAVGIFRAAVCTRG-----VAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 25.3%; Score 258; DB 2; Lu
Best Local Similarity 31.4%; Pred. No. 2.6c-15;
Matches 64; Conservative 34; Mismatches 84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25.5%; Score 260; DB 2; 28.9%; Pred. No. 1.7e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1132 VLHRGVKVTGVRYVKPWETLPKDS 1155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     171 VCTRGV-----AKAVDFIPVES 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 28.9%
les 65; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              硆
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
hypothetical protein homolog lin2744 [imported] - Listeria innocua (strain Clip11262) (Species: Listeria innocua (Cisteria innocua Cisteria innocua (Species: Listeria innocua Cisteria innocua (Cisteria innocua Cisteria innocua Cisteria innocua (Cisteria innocua Cisteria Cisteria Innocua Cisteria Comparative genomics of Listeria Species. Asteleccio number: Abilisteria Species Cisteria Cisteria Comparative genomics of Listeria Species.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Cross-references: GB:AL592022: PIDN:CAC97970.1; PID:g16415280; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
                                                                                                                                                                                                                                                                                                                                              113 VLACVPGTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT----LVGGTSKPFWLVPGGTG 168
                                                                                                                                                                                                                                                                                                                                                                                                                                     SLTPCTC----GSSDLYLVTRHADVIP-----VRRRGDSRGSLLSPRPISYLK--- 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.69 NNGNCGCGGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGGSSSSSSCIHIKVEN 228
                                                                                                                                                                    63 KAGSKLYATNGRL-----WEKELNGTGSWORVSSSSVPTDSDK-----KVMSIATDGNTF 112
                                                                                                                                                                                                                                                            57 LATCI -- NGVCWTVYHGAG -- - TRIIASPKGPVIQMYTNVDKDLVG -- -- - WPAPQGSR 105
                                                                                       3 KKGSVVIV--GRIVLNGAYAQQIRGEEGCQEISQ----IGRDKNQVEGEVQIVSTAAQIF 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKG---SVVIVGRIVLNGAYAQQTRGEBGCQETSQTGRDKNQVEGEVQIVSTAAQTFLAT
    71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1334;
78; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----GSSGGPLLCPAGHAVG 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 229 IDEQFLDMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGG 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 8.4%; Score 85.5; DB 2; I
Best Local Similarity 27.0%; Pred. No. 11;
Matches 58; Conservative 19; Mismatches 83;
```

	Search completed: August 30, 2003, 19:20:26 Job time : 18.0488 secs	Searc Job t
	902 YLTKAERGGIJTRADKARFIHRQLVETRQIJKNV 934	qa
	147 YL-KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAV 180	٥'n
	847 OSFITDNSIDNLVLISSAGNREKGDDVPPLEIVRRRKVFWEKLYGGNLMSKRRFD 901	qa
	102 QGSRSLTPCTCGSSDLYLVIRHADVIPVRRRGDSRGSLISPRPIS 146	δλ
		qa
_	60 CIVGVCWTVYHGAGTRIASPKGPVIOMYTNVDKDLVGWPAP 101	QY
	738 KKGILQSLKIVDELVSVMGYPPQTIVVEMARENQTTGKGKNNSRPRYKSLEKAIKEFGSQ 797	Q O

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 9.65768 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-12 1021 1 MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETTMRSP 195 Title: Perfect score: Sequence:

127863 seqs, 47026705 residues Searched:

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		ď			SUMMARIES	
Result No.	Score	Query Match	Leng	90	ID	iptio
7	894.5	7.	(*) !	: -	POLG_HCV1	P26664 h genome po
7	879.5	86.1	3011		POLG_HCVH	ъ о
٣	847	w.	(*)		POLG_HCVTW	h genome
4	842	ď	(*)		POLG_HCVBK	h genome
5	838	ς.	(*)	~	POLG_HCVJA	h genome
9	837	ď.	(*)		POLG_HCVJT	h genome
7	670.5	Š	(*)	•	POLG_HCVJ6	Ŀ
80	9	S.	(*)		POLG_HCVJ8	h genome
6	85.5			•	Y136_TREPA	treponema
10	82.5	•		٦	HHOA_ARATH	arabi
11	85	•			AAMP_HUMAN	Q13685 homo sapien
12	80.5				DEG1_ARATH	arabi
13	79				IBP1_HUMAN	
14	78.5	7.7			PAAD_PSEAE	pseudom
15	77.5	•			ENV_FLVGL	
16	77.5	•		•	CENE_HUMAN	
17	75.5	•		-	PO58_CAEEL	caenc
18	75				ENV_FLVSA	
19	74.5	٠			ZP3_RABIT	
20	74.5	7.3		-	HYPF_AZOVI	
21	74	٠			RAGE_RAT	Q63495 rattus norv
22	~	٠			ENV_RSVP	P03396 rous sarcom
23	73.5	•			GRAK_MOUSE	
24	m.	•			ENV_FLVCS	Q02077 feline leuk
25	Э.	•			I NVB_DAUCA	P80065 daucus caro
56		•	•		POL_GALV	P21414 gibbon ape
27	73	•			CAC3_BOVIN	Д
58	73	•			ENV_FSVSM	
58	73	•			ENV_MI,VFP	P26803 friend muri
30	73	٠			PTPO_MOUSE	P70289 mus musculu
31	7				POLC_LANVT	l gen
32	72.5	7.1			GRAM_HUMAN	α.
S) E)	7.7	7.1	629	7	VST2_HEVME	Q03500 hepatitis e

P26804 friend muri P39061 mus musculu P11033 mus musculu 090627 gallus gall P10705 herpes simp P97608 rattus norv P0321 epstein-bar P11261 feline leuk P54748 rattus norv P1261 bovine herp
ENV_MLVFF CALH_MOUSE GRAD_MOUSE GRAD_MOUSE UL21_HSV11 UL21_HSV11 UL21_HSV1E UL16_EBV ENV_FLVLB CN4A_RAT VGLB_HSVB2 ITA1_RAT
аналапалена
676 1527 248 248 535 1288 1336 662 844 917
110000000000
72 71.5 71.5 71.5 71.5 71.5 71.5
44444444444444444444444444444444444444

ALIGNMENTS

RESULT 1 PPOLG_HGV1 DO 1-A DO 1-A
--

```
1005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                  01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-S-EBP-2003 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.99-.); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P72); Nonstructural protein NS58 (P56); Nonstructural protein
NS58 (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-97331322; PubMed-9187654; Yao N. Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.; Structure of the hepatitis C virus RNA helicase domain."; Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.,
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
                                           52 AAQTFLATCINGVÇWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                  112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Structure 6:89-100(1998).
-!- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.
-!- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-!- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -:- FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-:- FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Prince A.M.;

"Genomic structure of the human prototype strain H of
"renomic structure of the human prototype strain H of
virus: comparison with American and Japanese isolates.

Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
3011 AA
                                                                                                                                                                         172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    of unwinding.
                                                                                                                                                                                                                                                             POLG_HCVH
                                                                                                                                                                                                                              POLG_HCVH
                                                                                                                                                                                                                                                                d
                                                                                                  ò
                                                                                                                             a
                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR RNVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS24 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                              Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87.6%; Score 894.5; DB 1; Length 3011;
85.8%; Pred. No. 2.2e-76;
Live 8; Mismatches 10; Indels 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GI.CNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                 rPro: IPR002868; NCV_NSSA.
rPro: IPR002166; HCV_RGRP.
rPro: IPR001659; Helicase_C.
rPro: IPR070995; RNA_pol_bS_PS.
rPro: IPR007094; RNA_pol_PSvir.
PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  327197 MW;
                                                                                                                                                                             Pfam; PF01542; ICV_core; 1. Pfam; PF01504; ICV_NS3; 1. Pfam; PF01004; ICV_NS4; 1. Pfam; PF01004; ICV_NS4; 1. Pfam; PF00201; ICV_NS5a; 1. Pfam; PF00994; Viral_RGRP; 1. Probom; PD186062; ICV_NS1; 1. Probom; PD186062; ICV_NS1; 1. PR0MPT; SM00487; DEXDC; 1.
                                                       HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2240
2364
2789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1615
1862
2013
3011
                          IPR002531;
IPR002518;
                                                                                                                                            InterPro; IPR007095;
                                                                                    IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                        IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity
Matches 175; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                    InterPro; InterPro;
                                                         InterPro;
InterPro;
                           InterPro;
InterPro;
                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
```

hepatitis

SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA. (RNA) (N

ď

Indels 11; Gaps

51

----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

KKGSVVIVG - - - RIVLNG - - -Conservative

```
(POTENTIAL)
        (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 879.5; DB 1;
Pred. No. 5.9e-75;
    (GLCNAC.
    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86.1%;
83.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            597
.598
.1611
1618
1623
1627
1636
      1411
1417
1420
1439
1463
1463
1478
1488
11490
11501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1333117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

133117

13311
      TURN
STRAND
HELIX
HELIX
THELIX
THELIX
THERN
STRAND
HELIX
TURN
HELIX
TURN
HELIX
TURN
HELIX
TURN
HELIX
TURN
HELIX
STRAND
                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
STRAND
TURN
HELLIX
HELLIX
HELLIX
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
TURN
STRAND
    CARBOHYD
CARBOHYD
                                                            CARBOHYD
CARBOHYD
                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best
      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (Sec http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                               | PERSI, MG7465; AAA45534.1; -
| PERSI, MG7465; AAA45534.1; -
| PERSI, MG7465; AAA45534.1; -
| PERSI, A181; 17-UNY-98.
| PERSI, 1A18; 17-UNY-98.
| PERSI, 1A18; 17-UNY-98.
| PERSI, A181; 17-UNY-98.
| PERSOPS: 829.001; -
| PERSOPS: 829.001; -
| PERSOPS: 829.001; -
| PERSOPS: 829.001; -
| PERSOPS: 199.001; -
|
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.
SIMILARITY: THE MS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE MS3 PROTEASE BELONGS TO PEPTIDASE FAMILY 229.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BY
(BY
(3Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1
CHARGE RELAY SYSTEM (1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      192
384
747
747
1022
1023
1105
1106
1116
1116
1116
1116
1230
1316
1316
1316
1316
1316
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CCHAIN
CCARBOHYD
CCARBOH
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                 1005 RRGQEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCT:TSLJTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROBUCE FROM N.A.

REDILINE-92230206; PubMed-1314449;

MEDILINE-92230206; PubMed-1314449;

Chen P.J. Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J. Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J. Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J. Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Chen P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Children P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Children P.J., Lin M.H., Tai K.F., Liu P.C., Lin C.J., Chen D.S.;

Children P.J., Lin M.H., Tai K.F., Lin P.C., Lin C.J., Chen D.S.;

Children P.J., Lin M.H., Tai K.F., Lin P.C., Lin C.J., Chen D.S.;

Children P.J., Lin M.H., Tai K.F., Chen D.S.;

Children P.J., Chen L. Lin M.H., Tai K.F., Chen D.S.;

Children P.J., Chen L. Lin M.H., Tai K.F., Chen D.S.;

Children P.J., Chen L. Lin M.H., Chen P.C., Chen D.S.;

Children P.J., Chen L. Lin M.H., Chen P.C., Chen D.S.;

Children P.J., Chen L. Lin M.H., Chen P.C., Chen D.S.;

Children P.J., Chen L. Lin M.H., Chen L. Lin M.H.,
                                                                                                                                                                                               52 AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                        112 CGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
66nome polyprotein (Contains: Corsid protein C (Core protein) (P22);
670-68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (GP8) (GP70) (NSI); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.22 -); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NSAB (P50) (NA-directed NA Polymerase) (EC 2.7.7.48)].
NS5B (P60) (P70) (RNA-directed NA Polymerase) (EC 2.7.7.48)].
Hepathis C virus (1solate Taiwan) (HCV).
                                                               3 KKGSVVIVG---RIVLNG----AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RNA)(N).
SUBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
   13; Indels
   9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1185 CTRGVAKAVDFIPVENLETTMRSF 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                    172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; M84754; -; NOT_ANNOTATED_CDS
PIR; A40244; GNWVIW.
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; $29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID-31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THE STATE OF STATE AND STA
                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
```

```
R Interpro; IPR007095; RNA_POL_DS_PS.

R Interpro; IPR007095; RNA_POL_DS_PS.

R Pfam; PF01542; HCV_core; 1.

R Pfam; PF01539; HCV_core; 1.

R Pfam; PF01539; HCV_NS1; 1.

R Pfam; PF01560; HCV_NS2; 1.

R Pfam; PF01001; HCV_NS2; 1.

R Pfam; PF01001; HCV_NS4; 1.

R Pfam; PF01001; HCV_NS4; 1.

R Pfam; PF01005; HCV_NS4; 1.

R Pfam; PF01005; HCV_NS4; 1.

R Pfam; PF00271; helicase.; 1.

R Pfam; PF00271; helicase.; NS1; 1.

R Pfam; PF00871; DEXDC; 1.

R Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS. DEC (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS. (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC. . .) (POTENTIAL) (GLCNAC. . .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC...) (F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  83.0%; Score 847; DB 1;
88.2%; Pred. No. 7.1e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
HCV_core.
HCV_env.
HCV_NSI.
HCV_NS2.
HCV_NS3.
HCV_NS4.
HCV_NS4B.
HCV_NS5B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1083
                                                                                               IPR000745;
                                                                                                                   InterPro; IPR001490;
                                                                                                                                     InterPro; IPR002868;
                                                                                                                                                         IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity
Matches 157; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure.
INIT_MET
                                                                                                                                                         nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
```

ö

Gaps

;

Indels

Mismatches

11;

Conservative

InterPro; IPR001410; DEAD.

Page

us-09-965-594-12.rsp

PROTEIN C AND MRNA

```
1031 AYAQQTRG1.FGCITTSL.FGRDKNQVEGEVQVVSTATQSF1.ATCINGVCWTVYHGAGSKTL 1090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            X *RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1679-1691.

X MEDLINE-98227846; PubMed-9568891;
Yan Y., Li Y., Munshi S., Sardana W., Sardana M., Steinkuehler C., Tomei L., de Francesco R., Kuo I.C., Chen Z.;
T *Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998).

C *Protein Sci. 7:837-847(1998).

C *Protein Sci. 7:837-847(1998).

C *Protein Sci. 7:837-847(1998).

C *PROTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PROLE IN THE VIRAL RNA REPLICATION.

C *PROTICT ACTIVITY: Hydrolysis of four peptide bonds in the viral presuresor polyprotein, commonly with Asp or Glu in the P6 precursor polyprotein.

C *PRALYTIC ACTIVITY: N nucleoside triphosphate ~ N diphosphate +
                                                                                                                                             ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRKGDSRG 137
AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                     SILSPRPISYLKGSSGGPLICPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-5EP-2003 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; NonStructural protein NS2 (F2 3 4 .22 . .); Protease/halacase NS3 (P70) (Hepacivirin)
(EC 3 4 .21 .98); Nonstructural protein NS4A (P4); Nonstructural protein NS5B (P65) (P70); (NNA-directed NNA POlymerase) (EC 2.7.7.48)].
NS5B (P65) (P70) (SNA-directed NNA Polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate BK) (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Takamizawa A., Mori C., Fukc I., Manabo S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.

BUDILNE-PRO1508B; Pubmed-8861916;

LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N., Moomaw E.W., Adachi T., Hostomska Z.;

The crystal structure of hepatitis C virus NS3 proteinase reveals trypsin-like fold and a structural zinc binding site.";

Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-96235224; PubMed-8647104;
Borcowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Riochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A. MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1487-1500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                              138
                                                                                                                                                 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 4
POLG_HCVBK
ID POLGAR
ID POLGAR
ID POLGAR
ID POLGAR
ID POLGAR
ID OLANG
DT 01-AUG
DT 01-AUG
DE GENOME
DE GENOME
DE GENOME
DE GENOME
DE GENOME
ON NASB (
ON NASB
                                                          q
                                                                                                                                             ò
                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                q
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EWBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E) (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.1.PZ (POTENTIAL).
PRONTRUCTURAL PROTEIN NS.2 (POTENTIAL).
PROTEASE/HELICASE NS.3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS.4 (POTENTIAL).
RIMA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART: SM00487; DEXDC; 1.
Subjectedin; Glycoptedan; Transferase; RNA-directed RNA polymerase; Core protein; Glycoptedin; Halicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001410; DEAD.
InterPro; IPR001212; HCv_capsid.
InterPro; IPR001212; HCv_capsid.
InterPro; IPR001213; HCv_care.
InterPro; IPR001219; HCv_care.
InterPro; IPR001219; HCv_MS1.
InterPro; IPR0014109; HCv_MS3.
InterPro; IPR001409; HCv_MS4.
InterPro; IPR001409; HCv_MS4.
InterPro; IPR001409; HCv_MS4.
InterPro; IPR001868; HCv_MS5a.
InterPro; IPR001868; HCv_MS5a.
InterPro; IPR001969; HCv_MS1.
InterPro; IPR001969; HCv_MS1.
InterPro; IPR001969; HCv_MS1.
InterPro; IPR001999; RNA_pol_DSy1.
InterPro; IPR001099; RNA_pol_DSy1.
InterPro; IPR001999; RNA_pol_DSy1.
IPRam; PF01542; HCv_Core; IPRam; PF01538; HCv_MS1; IPRam; PF01538; HCv_MS1; IPRam; PF01538; HCv_MS1; IPRam; PF01509; HCv_MS1; IPRam; PF01509; HCv_MS3; IPRAm; PF01001; HCv_MS4b; IPRAm; PF01001; HCv_MS4b; IPRAm; PF01001; HCv_MS4b; IPRAm; PF01601; HCV_MS4b; IPRAm;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                       EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1615
1862
2013
3010
369
1083
11065
11165
1234
209
                                                                                                                                                                                                                                                           1A1Q; 25-MAR-98.
1JXP; 14-JAN-98.
1DXP; 16-APK-98.
1C2P; 15-NOV-00.
1CSJ; 08-NOV-99.
1GX5; 09-APR-02.
1GX6; 10-APR-02.
1GX6; 10-APR-02.
                                                                                                                                                                                                                                            GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1165
1230
1316
196
209
234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 347
1083
1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; S29.00
MEROPS; U39.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
ENII_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                               PDB;
PDB;
PDB;
PDB;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  78 ASPKGPVIQMYTHVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWIVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFKAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82.5%; Score 842; DB 1; Length 3010;
86.5%; Pred. No. 2.1e-71;
iive 14; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                   (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                         (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                  (GLCNAC.
(GLCNAC.
                                              (GLCNAC.
                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3010 AA
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVJA STANDARD;
P26662;
01-AUG-1992 (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 86.5%
Matches 154; Conservative
                                                                                                                                                                                                                                                                       2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 2004 | 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11188
11197
1202
1204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ą.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1203
1680
3010
                                                                                                                                                                                                                                                                                                                                                                                         100334
100334
100334
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
10036
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100336
100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1151
                     CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                              CARBOHYD
                                                                                                                 CARBOHYD
                                                                                                                                     CARBOHYD
                                                                                                                                                               CARBOHYD
                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                          HELIX
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TURN
STRAND
HELIX
TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
STRAND
HELIX
TORN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLG_HCVJA
ID POLG_HC
AC P26662
DT 01-AUG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

```
12 - NGC-1907 (Rel. 12, Last sequence update)

26 - Readone polypotreen (Contains: Capsed protein) (22);

27 - Readone polypotreen (Contains: Capsed protein) (22);

28 - Readone polypotreen (Contains: Capsed protein) (22);

29 - Readone polypotreen (Contains: Capsed protein) (22);

20 - Readone polypotreen (Contains: Capsed protein) (22);

20 - Readone polypotreen (Contains: Capsed protein) (22);

20 - Readone polypotreen (Contains: Capsed protein) (22);

21 - Readone polypotreen (Contains: Capsed protein) (22);

22 - Readone polypotreen (Contains: Capsed protein) (22);

23 - Readone polypotreen (Contains: Capsed protein) (22);

24 - Readone polypotreen (Contains: Capsed protein) (22);

25 - Readone polypotreen (Contains: Capsed protein) (22);

26 - Readone polypotreen (Contains: Capsed protein) (22);

27 - Readone polypotreen (Contains: Capsed protein) (22);

28 - Readone polypotreen (Contains: Capsed protein) (22);

29 - Readone polypotreen (Contains: Capsed protein) (22);

20 - Readone polypotreen (Contains: Capsed protein) (23);

27 - Readone polypotreen (Contains: Capsed protein) (23);

28 - Readone polypotreen (Contains: Capsed protein) (23);

29 - Readone polypotreen (Contains: Capsed protein) (23);

29 - Readone polypotreen (Contains: Capsed protein) (23);

20 - Readone polypotreen (Contains: Capsed protein) (23);

21 - Readone polypotreen (Contains: Capsed protein) (23);

22 - Readone polypotreen (Contains: Capsed protein) (23);

23 - Readone polypotreen (Contains: Capsed protein) (24);

24 - Readone polypotreen (Contains: Capsed protein) (24);

25 - Readone polypotreen (Contains: Capsed protein) (24);

26 - Readone polypotreen (Contains: Capsed protein) (24);

27 - Readone polypotreen (Contains: Capsed protein) (24);

28 - Readone polypotreen (24);

28 - Readone polypot
```

```
MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRGDSRG 137
Pfam; PF01538; HCV_NS2: 1.

R Pfam; PF01006; HCV_NS4: 1.

R Pfam; PF01006; HCV_NS4: 1.

R Pfam; PF01006; HCV_NS4: 1.

R Pfam; PF01001; HCV_NS4: 1.

R Pfam; PF01506; HCV_NS5: 1.

R Pfam; PF00271; Helicase. 2: 1.

R Pfam; PF00291; Niral_RdRP; 1.

R Probom; PD186062; HCV_NS1; 1.

R RAPR; DEXDC: 1.

R Probom; PD186062; HCV_NS1; 1.

R Probom; PD18606; HCV_NS2; HCRITIAL).

R Probom; PD18606; HCV_NS1; 1.

R Probom; PD18606; HCV_NS1; HCRITIAL).

R CHAIN 1007 1616; HCV_NS2; HCRITIAL).

R PROBOP R PROBOM; PROFINIAL PROTEIN NS48 (POTENTIAL).

R PROBOP R PROBOM; PROFINIAL PROFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTF:LATCINGVCWTVYHGAGTRT1 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SLLSPRPISYLKGSSGGPLLCPAGHAVGIFKAAVCTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHANGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
DECH POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 838; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          85.4%; Pred. No. 5.1e-71;
tive 16; Mismatches 10; Indeis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED GLCNAC
N-LINKED GLCNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2041 2041 N-LIN
2077 2077 N-LIN
2240 2240 N-LIN
2788 2788 N-LIN
3010 AA: 327017 MW; A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 85.4 Matches 152; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Rel. 25, (Rel. 25, I (Rel. 25, I (Rel. 42, I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 6
POLG_HCVJT
ID POLG HCVJT
AC Q00269;
DT 01-APR-1993 (
DT 01-APR-1993 (
DT 15-SEP-2003 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYL
                 Sort State S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED custatation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                   MEDLINE-92295714; PubMed-1318627;
A Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Anakazwa T., Hijikata W., Ishimura Y., Shimotohno K.;
Nakazwa T., Hijikata W., Ishimura Y., Shimotohno K.;
Molecular cloning of hepatitis C virus genome from a single Japanese carier: sequence variation within the same individual and among infected individuals.";
Virus Res. 23:39-53(1992).
L. PUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE
I. FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4B AND NSS AND NSS MAY PROLES HOUSE IN THE VIRAL RNA REPLICATION.
NSS AND NSS MAY HYDROLYSIS of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thi in Pl and Ser or Ala in Pl'.
C. CATALYTIC ACTIVITY: N nucleoside triphosphate * N diphosphate +
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (GR01) (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -1. SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E, THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro: IPR001410; DEAD.
InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_core.
InterPro: IPR002519; HCV_env.
InterPro: IPR002531; HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D11168; BAA01943.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       helicase_C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF00998; Vir
i; PD186062;
                                                                                                                                                                                                                                                                                      NCBI_TaxID=31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                039.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (RNA)(N).
                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF(
ProDom; 1
```

Ø

```
1031 AYAQQTRGILGCIVTSLTGRDKNQVEGEVQVVSTATQSFLATCVNGVCWTVFHGAGSKTL 1090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYI.VTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                        11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELULAR AMINOPEPIDASE.
CARSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELORE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS!/F2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2013 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35): Envelope glycoprotein E2
(GP68) (GP70) (NS1): Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-): Protease/Relicase NS3 (P70) (Bepaciation)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                        AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                         RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
SMART; SM00487; DEXDc; 1.
Subsprotelar, Glycoprotelar, Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding: Transmembrane; Nonstructural protein; Hydrolase; Serine protese;
                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                              REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
                                                                                                                                                                                                                                              (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                            (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                     11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                 94A1C77435D642BB CRC64;
                                                                                                                                                           CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                  82.0%; Score 837; DB 1;
86.5%; Pred. No. 6.3e-71;
                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                  (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                   GLCNAC.
                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                           GLCNAC
                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                         ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                     13; Mismatches
                                                                                                                                                                                                          N.LINKED
                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                 326573
                                                                                                                                                                                                                                                                                                                                                                                                                                    154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                  ¥
                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                  2041
2077
2240
2529
2788
3010
                                                                                                                                                           3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 138
                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78
                                                                                                                                                                                                                                                                                                                                                                                                                                                        18
                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                              SITE
                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                    TRANSMEM
                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                         NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 7
POLG HCWJ6
LD POLGCD H
AC P26660
DT 01-AUG
DT 21-AUG
DT 28-PEB
DE Genome
DE ENVELO
DE (GP68)
DE (GP68)
DE (EC 3.
                                                                                                     CHAIN
CHAIN
CHAIN
                                                                          CHAIN
                                                                                                                                  CHAIN
                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loc
Matches
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                       MEDLINE-92044440; PubMed-1658196; Okamoro H., Okada S.-I., Suglyama Y., Kurai K., Lizuka H., Okamoro H., Mayumi M.; Machida A., Miyakawa Y., Mayumi M.; Machida A., Miyakawa Y. Mayumi M.; "Nucleotlide sequence of the genomic RNA of hepatitis C virus isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -i-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPOTRIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine proteises. INTL-MET.

INTL-MET.
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (1solate HC-J6) (HCV).
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .. SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ...v_RdRP.
...vo50; Helicase.C.
...pR007095; RNA_Dol_DS_PS.
); IPR007094; RNA_Dol_DS_PS.
); IPR34; HCV_cappsid; 1.
1542; HCV_core; 1.
539; HCV_core; 1.
60; HCV_NS; 1.
'; HCV_NS; 1.
'; HCV_NS; 1.
'; HCV_NS; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4a; 1.
HCV_NS4b; 1.
HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; JQ1303; JQ1303.
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001490;
                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                        NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF01006;
Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            039.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (RNA)(N).
                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
```

```
THEI.
                  NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PF01543;
Pfam; PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HSSP; P27958;
                                                                                                                                                                                                                                                                                             {RNA}(N).
                                                                                                                                                  genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART;
   THE HELL WAS A PRESENT TO BE A PRESENT TO BE A PRESENT TO COUCCO 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1010 IGREVILGPADGYTSKGWSLLAPITAYAQQTRGLLGTIVVSMTGRDKTEOAGEIOVI,STV 1069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             113 GSSDLYLVTRHADVIPVRHRCDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVC 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24). Last annotation update)
03-REB-2003 (Rel. 41, Last annotation update)
04-One polyprotein [Contains: Capsid protein C (Core protein) (P22);
05-05 (GP68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (P20)
05-05 (GP70) (NSI); Protease/helicase NS3 (P70) (Hepacivirin)
05-05 (P27); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (NAN-directed RNA polymerase) (EC 2.7.7.48)].
05-05 (P66) (P70) (NAN-directed RNA polymerase) (EC 2.7.7.48)].
05-05 (P65) (P70) (NAN-directed NA polymerase) (EC 2.7.7.48)].
                            MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS! (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
RONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10 VGRIVLNG -------AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17; Gaps
                                                                                                                                                                           CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 670.5; DB 1; Length 3033; Pred. No. 3.8c-55;
                CAPSID PROTEIN C (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F957F5C1A273BE9E CRC64;
 CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                         (GICNAC
                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                          ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3033 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27; Mismatches
                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                           N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            173 TRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65.7%;
62.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      649
1091
2038
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               126;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJ8
P26661;
                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                            ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                               FRANSMEN
                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
POLG_HCVJ8
                                                                                                              CHAIN
                                                               CHAIN
                                                                                                                               CHAIN
                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
```

```
the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                      Okamoto H., Kurai K., Okada S. I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S., "Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                            in the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -! - SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
CHAIN 192 383 MAJOR ENVELOPE PROTEIN E (POTENTIAL).
CHAIN 384 733 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
CHAIN 734 1010 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                 VITOLOGY 188:331-341(1992).

-1- FUNCTION: THE SMALL PROTEINS NSZA, NSZB, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-1- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Gys or Thr in P1 and Ser or Ala in P1'.

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | IPR000745; HCV_NS4a | IPR001490; HCV_NS4b | IPR001490; HCV_NS4b | IPR001266; HCV_RGAP | IPR002166; HCV_RGAP | IPR007095; RNA_POL_PS-PS | IPR007094; RNA_POL_PS-PS | IPROV_COTE: 1.
SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01538; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01566; HCV_NS5s; 1.
Pfam; PF001569; Viral_RGRP; 1.
ProDom; PP0186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002531; HCV_NS1
IPR002518; HCV_NS2
IPR004109; HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; Di
InterPro; IPR002522; H
InterPro; IPR002521; H
InterPro; IPR002519; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SM00487; DEXDC;
```

```
48984 MW;
                                                                 spirochete.";
Science 281:375-388(1998),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                               23
485
24
178
210
267
327
                                                                                                                                                                                                                                                                                                                                                               24
24
164
196
253
318
444
485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1010 LGREVLLGPADGYTSKGWKLLAPITAYTQQTRGLLGAIVVSLIGRDKNEQAGQVQVLSSV 1069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   53 AQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22
                       RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1070 TQTFLGTSISGVLWTVYHGAGNKTLACPKGPVTQMYTSAEGDLVGWPSPPGTKSLDPCTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10 VGRIVLNG -------AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Odoson R., Gwinn K.A., Sodorgon E., Gwinn K.A., Sodorgren E., Hardham J.M., McLood M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
               NONSTRUCTURAL PROTEIN NS4A (POTENTIAL)
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Treponema pallidum.
Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
NCBL_TaxID=160;
                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                (POTENTIAL).
                                                                                                                                                                                                     (POTENTIAL)
 PROTEASE/HELICASE NS3 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                 (GLCNAC...) (GLCNAC...) (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65.5%; Score 668.5; DB 1; 62.6%; Pred. No. 5.9e-55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25; Mismatches 34;
                                                                                                                                                                                                       (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                 (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                        N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              173 TRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16-OCT-2001 (Rel. 40, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 127; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                            649
1091
2038
2359
2811
                                                                                                                                                                                                                                                                                                                                                                                          649 6
1091 10
2038 20
2359 23
2811 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Y136_TREPA
083172;
                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                           ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SECUENCE
                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local
                                                               TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     rP0136.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-slb.ch/announce/or send an email to license@isb-slb.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53 KAGSKLYATNGRL-----WEKELNGTGSWQKVSSSSVPTDSDK-----KVMSIATDGNTF 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               106 SLTPCTC----GSSDLYLVTRHADVIP-----VRRRGDSRGSLLSPRPISYLK--- 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          57 LATCI -- NGVCWTVYHGAG---TRIIASPKGPVIQMYINVDKDLVG------WPAPQGSR 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             159 NNGNCGCGGGGGSSSSSSCIHIMLVPGGTGNNGNCGCGGGGGGSSSSSSSCIHIKVEN 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KRGSVVIV--GRIVLNGAYAQQTRGEEGCQETSQ----TGRDKNQVEGEVQIVSTAAQTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease,
a homolog to the bacterial periplasmic protease HhoA.";
Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09SEL7: 049507;
16-007-2001 (Rel. 40, Created)
16-007-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HhoA, chloroplast precursor (EC 3.4.21.-),
Arabidopsis thaliana (Mouse-ear cress),
Arabidopsis thaliana (Mouse-ear cress),
Bukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnollophyta; eudicotyledons; core eudicots; Rosidae;
curosids II; Brassicales; Brassicaceae; Arabidopsis.
McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.
Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
Venter J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71; Gaps
                                                                                                                                                                               :- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
                                                                                            Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HYPOTHETICAL LIPOPROTEIN TP0136.
N-ACYL DIGLYCERIDE (POTENTIAL).
GLY/SER-RICH.
                                                                                                                                                                                                                 (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.4%; Score 85.5; DB 1; Length 485; 33.6%; Pred. No. 1.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----GSSGGPLLCPAGHAVG 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               219 IDEQFLDMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGGG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 23.6%; Preu. ....
Matches 52; Conservative 19; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLY/SER-RICH.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLY/SER-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AE001199; AAC65137.1; ALT_INIT.
TIGR; TP0136; -.
Hypothetical protein; Lipoprotein; Mer
```

.. 8

Gaps

39;

Indels

,64

26.2%; Pred. no.+ive 22; Mismatches

Conservative

Best Local Similarity Matches 39; Conserv

73 GTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRR 132

8.1%; Score 82.5; DB 1; Length 321;

34691 MW;

321 AA;

SEQUENCE Query Match

DOMAIN ACT_SITE ACT_SITE ACT_SITE CONFLICT

STATE TELLICAN SOLUTION SOLUTI

Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide.
TRANSIT 1 26 CHLOROPLAST (POTENTIAL).
TRANSIT 27 71 THYLAKOID.
PROTEASE HHOA.

MEROPS; SO1.279; InterPro; IPRO01940; Protease2C. InterPro; IPRO01254; Ser_protease_Try. Pfam; PF00089; trypsin; I. PRINTS; PRO0894; PROTEASES2C.

CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
R -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;

---SLLSPRPISYLK-----GSSGGPLLCPAGHAVGIF 167

133 GDSRG------

g ò 셤

220 GNPYGYENTLIIGVVSGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSYGHTIGV- 278

168 RAAVCTR---GVAKAVDF-IPVESLETTM 192

```
Nature 402:769-777(1999).
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its mose by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                      Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                    15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
279 NTATFTRKGSGMSSGVNFAIPIDTVVRTV 307
                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                               MEDL, INE = 95262124; PubMed = 7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M95627; AAA68889.1; -. PIR; I39383; I39383.
                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genew; HGNC:18; AAMP.
MIM; 603488; -.
                                                                                                                                                                                                           NCBI_TaxID=9606;
                                                                                                   15-JUL-1998 (15-JUL-1998 (28-FEB-2003 (
                                                                                                                                                                                                                                                     TISSUE-Brain;
                                                                        AAMP_HUMAN
Q13685;
                                                                AAMP_HUMAN
                                                                              g
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

EMBL; AF114386; AAF24060.1; '. EMBL; AL021710; CAA16717.1; ALT_SEO. EMBL; AL161548; CAB78839.1; ALT_SEO.

```
Hydrolase; S
TRANSIT
TRANSII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IBPL_HUMAN
ID IBPL_H(
AC P08833
DT 01-NOV
                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                      66 WTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSL-----TPCTCGSSDLYLV 120
                                                                                                                                                                                                                                                                                    121 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA------160
                                                                                                                                                                                                                                                                                                                                       241 -----VLPDGKRAVVGYEDGTIRIWDLKQGSP1HVLKGTEGHQGPLTCVAANQDGSL1LT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                   DEGL_ARATH STANDARD; PRT; 437 AA.
022609; O9LK85;
16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
28-FBB-2003 (Rel. 41, Last annotation update)
Protecase Do-like 1, chloroplast precursor (EC 3.4.21.-).
Protecase Do-like 1, Tast annotation update)
Arabidopsis thaliana (Mouse-ear cress).
Arabidopsis thaliana (Mouse-ear cress).
Bukaryota: Viridiplantae; Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Wagnoliophyta: cudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceac; Arabidopsis.
                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-CV. COLUMDIA:
MEDLINE-20363099; PubMed-10907853;
REMENO T., KATOH T., SATO S., NAKAMURA A., ASAMIZU E., Tabata S.;
KANEKO T., KATOH T., SATO S., SAKAMURA A., ASAMIZU E., Tabata S.;
SETUCLURAL ANALYSIS OF ARabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-cv. Columbia:
Kaseslbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh I., Lindahl M., Cook M., Adam Z.;
Ildentification and characterization of Degp, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                  94;
                                                                                                                                                                                                                                                                                                                                                            161 ------GHAVGIFR----AAVCTRGVAKAVDFIPVESL 188
                                                                                                                                                                                                                                                                                                                                                                        296 GSVDCQAKLVSATIGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                         8.0%; Score 82; DB 1; Length 452; 25.3%; Pred. No. 2.6;
                                                                                                                                                                                                                                                13; Mismatches 47; Indels
                                                                                           HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                     DA1413D25EB236C0 CRC64;
GO; GO:0008201; F:heparin binding activity; TAS.
InterPro; IPR001680; WD40.
Pfam: PF00400; WD40; 8.
SMART; SM00320; WD40; 8.
PROSITE; PS500678; WD_REPEATS_1; 1.
PROSITE; PS50082; WD_REPEATS_2; 6.
PROSITE; PS50294; WD_REPEATS_REGION; 1.
                                                                                                      POLY-GLU.
                                                                                                               49015 MW;
                                                                                                                                                                                                                                     Best Local Similarity 25.3%
Matches 42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA Res. 7:217-221(2000)
                                                                                          1138
177
1138
1180
220
220
261
305
404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [3]
SEQUENCE OF 104-118.
                                                                                                                                                                                                     452 AA;
                                                                                 Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BAC clones.
                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                            Query Match
                                                                                                    DOMAIN
REPEAT
REPEAT
REPEAT
                                                                                                                                               REPEAT
REPEAT
                                                                                                                                                                               REPEAT
REPEAT
                                                                                           DOMAIN
                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
DEGLARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and
                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                            ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68 VYHGAGTRIIASPKGPVIQMY-----PA 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRRGDSRGSLLSPRPI 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                              CAN DEGRADE BETA-CASEIN.
ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Transit peptide; Chloroplast; Thylakoid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        146 SYL------KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF-IPVESL 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : |:||||| : | |||||| : | ODVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGASSGVGFSIPVDTV 317
                                                                                                                                     O-PHENANTHROLINE.
-1- SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.9%; Score 80.5; DB 1; Length 437; 25.6%; Pred. No. 3.5; tive 18; Mismatches 55; Indels 59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2).

V -> I (IN REF. 2).

P -> S (IN REF. 2).

G -> R (IN REF. 2).

LL -> D (IN REF. 2).

LL -> V (IN REF. 2).

I 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHLOROPLAST (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROTEASE DO-LIKE 1. SERINE PROTEASE.
                                                                                                                                                                                                                                                                                                           -:- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C-:- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001940; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00595; PD2; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00814; PROTEASES2C.
SMART; SM00228; PDZ; 1.
PROSITE; PS50106; PDZ; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IBP1_HUMAN STANDARD;
P08833;
Ol-NOV-1988 (Rel. 09, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46213 MW;
                                                                                                                                                                                                                                                                  -! - INDUCTION: By heat shock.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; S01.279; -. InterPro; IPR001478; PDZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                355
335
331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                321
421
171
201
280
23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36
54
66
64
68
355
381
416
437 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                 MEMBRANE.
```

```
SEQUENCE OF 26-53.
                                                                                                                                                                                                                                                                                                    MUTAGENESIS.
   "Insulin-like growth factor (IGF) binding protein complementary deoxyribonucleic acid from human HEP G2 hepatoma cells: predicted protein sequence suggests an IGF binding domain different from those of the IGF-1 and IGF-11 receptors.";
                                                                                                                SEQUENCE FROM N.A.
TISSUE-Placenta;
MEDLINE-89052654; PubMed-2461294;
Brinkman A., Groffen C., Kortleve D.J., Geurts A., Drop S.L.S.;
"Isolation and characterization of a cDNA encoding the low molecular weight insulin-like growth factor binding protein (IBP-1).";
EMBO J. 7:2417-2423(1988).
                                                                                                                                                                                                                                                                                                                                                        Grundmann U., Nerlich C., Bohn H., Rein T.;
"Cloning of cDNA encoding human placental protein 12 (PP12): binding protein for IGF I and somatomedin.";
Nucleic Acids Res. 16:8711(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Julkunen M., Koistinen R., Aalto-Setala K., Seppala M., Janne O.A.,
                                                                                                                                                                                                                                                 Brewer M.T., Stetler G.L., Squires C.H., Thompson R.C., Busby W.H. Jr., Clemmons D.R.; Cloning, characterization, and expression of a human insulin-like growth factor binding protein. 152:1289-1297(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. MEDIJNE-89087480; PubMed-2849945; MEDIJNE-89087480; PubMed-2849945; Medilyme-89087480; PubMed-2849945; Corganization of the gene encoding the insulin-like growth factor binding protein IBP-1.*; Biochem. Blophys. Res. Commun. 157:898-907(1988).
                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBL_TaxID=9606;
            28-FEB-2003 (Rel. 41, Last annotation update)
Insulin-like growth factor binding protein 1 precursor (IGFBP-1)
(IBP-1) (IGF-binding protein 1) (Placontal protein 12) (PP12).
IGFBP1 OR IBP1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kontula K.;
"Primary structure of human insulin-like growth factor-binding
protein/placental protein 12 and tissue-specific expression of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [6]
SEQUENCE FROM N.A.
MEDLINE-89330502; PubMed-2474129;
Cubbage M.L., Suwanichkul A., Powell D.R.;
"Structure of the human chromosomal gene for the 25 kilodalton insuln-like growth factor binding protein.";
Mol. Endocrinol. 3:846-851(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lee Y.-L., Hintz R.L., James P.M., Lee P.D.K., Shively J.E., Powell D.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-92217971; PubMed-1373120;
Ehrenborg E., Larsson C., Stern I., Janson M., Powell D.R.,
 01-NOV-1988 (Rel. 09, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          of the IGF-I and IGF-II receptors."; Mol. Endocrinol. 2:404-411(1988).
                                                                                                                                                                                                                                                                                                                                            MEDLINE-88335621; PubMed-3419931;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             [5]
SEQUENCE FROM N.A.
MEDLINE-88334540; PubMed-2458522;
                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-88312985; PubMed-2457513;
                                                                                                                                                                                                                                         MEDLINE-88240345; Pubmed-2454104;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEBS Lett. 236:295-302(1988).
                                                              Homo sapiens (Human).
                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                               TISSUE-Decidua;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mRNA.
```

```
neutron-orocous; rupwed-1032900;

Neumann G.M., Bach L.A.;

The N-terminal disulfde linkages of human insulin-like growth a The N-terminal disulfde linkages of human insulin-like growth factor-binding protein-6 (hIGFBP-6) and hIGFBP-1 are different as a determined by mass spectrometry.;

T determined by mass spectrometry.;

T Biol. Chem. 274:14587-14594 (1999).

-!- FUNCTION: IGF-binding proteins prolong the half-life of the IGFs and have been shown to either inhibit or stimulate the growth promoting effects of the IGFs on cell culture. They alter the interaction of IGFs with their cell surface receptors.

-!- PIM: Phosphorylated; probably by casein kinase II. Alters the affinity of the protein for IGFs.

-!- FIM: Phosphorylated; probably by casein kinase II. Alters the affinity of the protein for IGFs.

-!- SIMILARITY: Contains 1 IGFBP domain.

-!- SIMILARITY: Contains 1 IGFBP domain.

-!- SIMILARITY: Contains 1 IHyroglobulin type-1 domain.

-!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                       MEDLINE-8910723; PubMed-2466665; Luthman H., Soederling-Barros J., Persson B., Engberg C., Stern I., Luthman H., Soederling-Barros J., Persson B., Lindgren B., Lake M., Franzen S.A., Israelsson M., Raden B., Lindgren B., Hjelmqvist L., Encerbaeck S., Carlsson P., Bjursell G., Povoa G., Hall K., Joernvall H.; "Human insulin-like growth-factor-binding protein. Low-molecular-mass form: protein sequence and cDNA cloning."; Eur. J. Biochem. 180:259-265(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Increase of beta-actin mRNA upon hypotonic perfusion of perfused rat
*Contiguous localization of the genes encoding human insulin-like growth factor binding proteins 1 (IGBP1) and 3 (IGBP3) on chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Busby W.H. Jr., Klapper D.G., Clemmons D.R.; Purification of a 31,000-dalton insulin-like growth factor binding protein from human amnlotic fluid. Isolation of two forms with different biologic actions.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Identification of the sites of phosphorylation in insulin-like growth factor binding protein-1. Regulation of its affinity by phosphorylation of serine 101.";
J. Biol. Chem. 268:1125-1131(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-93123224; PubMed-7678248;
Jones J.I., Busby W.H. Jr., Wright G., Smith C.E., Kimack N.M.,
Clemmons D.R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92070504; PubMed-1959616;
Brinkman A., Kortlrve D.J., Schuller A.G.P., Zwarthoff E.C.,
Drop S.L.S.;
                                                                                                                                                                                               19)
SEQUENCE OF 141-259 FROM N.A., AND SEQUENCE OF 26-259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PHOSPHORYLATION SITES, AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Biol. Chem. 263:14203-14210(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; Y00856; CAA68770.1; -.
EMBL; M20841; AAA52540.1; ALT_FRAME.
EMBL; X12385; CAA30942.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISÚLFIDE BONDS.
MEDLINE-99262603; PubMed-10329650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-89008261; PubMed-2971653;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEBS Lett. 292:264-268(1991).
                                                                                                                                    Genomics 12:497-502(1992).
```

```
MEDLINE-86200439; PubMed-3009890;
                                                                                                                                                                                                                                                             EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182 FIPVESLET 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 FVVARILNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ENV_FLVGL
P08359;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 RVVESLAKAQETSGEEISKFYLPNCNKNGFYHSRQCETSMDGEA----- 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12 RIVLNGAYAQQTRGEE------GCOEIS--OIGRDKNQVEGEVQIVSTAAQTFLATCIN 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28; Gaps
                                                                                                                                                                   MIN: 146730; -
60. 60:0005615; C:extracellular space; TAS.
60. 60:0005520; F:insulin-like growth factor binding activity; TAS.
60. 60:0007165; P:signal transduction; TAS.
InterPro; IPR000867; Insl_gro_fac_pr.
InterPro; IPR000716; Thyrogiobulin_l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                INSULIN-LIKE GROWTH FACTOR BINDING
                                                                                                                                                                                                                                  Pfam: PF00219; IGFBP; 1.
SMART; SM00121; IB: 1.
SMART; SM00121; IB: 1.
SMART; SM00212; IGF BINDING; 1.
PROSITE; PS00484; THYROGLOBULIN_1; 1.
Growth factor binding; Signal; Phosphorylation; Polymorphism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.7%; Score 79; DB 1; Length 259; 28.3%; Pred. No. 2.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12; Mismatches 26; Indels
                                                                                                                                                                                                                                                                                                                                                                                                            PHOSPHORYLATION (MAJOR).
PHOSPHORYLATION.
PHOSPHORYLATION.
V -> I (IN dDSNP:1065782).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -> Q (IN REF. 2).
8AA75AF7DC966012 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         /FTId-VAR_011905.
I -> M (IN dbSNP:4619).
/FTId-VAR_003821.
                                                                                                                                                                                                                                                                                                                                          PROTEIN 1.
THYROGLOBULIN IYPE I.
CELL ATTACHMENT SITE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (RC 4.1.1.-)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63 GVCWTVYHGAGTRIIASPK---GPVIOMYTNV 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       209 AA
Pseudomonadaceae; Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27903 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity 28.3%
Matches 26; Conservative
                                                                                                                                ITRANSFAC: T00400;
Siena-2DPAGE: P08833; -.
Genew; HGNC:5469; IGFBP1.
MIM: 146730; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           213 2
259 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAAD_PSEAE
Q9HX08;
                                                                                                                                                                                                                                                                                                                                                                          DISULFID
DISULFID
DISULFID
MOD_RES
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARIANT
                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PA4019
                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAAD_PSEAE
```

g

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84 VIQMYTNVDKDLVGWPAPQGSRSLTP----CTCGSSDL-----YLVTRHADVIPV 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   75 OIRVFGOND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALK 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29 QEEREVHFLISKAAQLVMAT------ETDVALPAKPQAMQAFLTEYCGAAAG 74
STRAIN-ATCC 15692 / PAO1;
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Sicover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Elckey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            130 RRRGDSRGSLLSPR--PIS----YLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41 QVEGEVQ-IVSTAAQIFLATCINGVCWIVYHGAGIRIIASPKGP---------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Onions D.E., Neil J.C.;
"Nucleotide sequences of a feline leukemia virus subgroup A envelope gene and long terminal repeat and evidence for the recombinational origin of subgroup B viruses.";
J. Virol. 58:825-834(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR: H83144; H83144.
InterPro; IPR003382; Flavoprotein.
Pfam; PF02441; Flavoprotein; 1.
Hypothetical protein; Lyses: Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MW; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Stewart M.A., Warnock M., Wheeler A., Wilkie N., Mullins J.I.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nature 406:959-964(2000).
-:- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE / PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-AUG-1988 (Rel. 08, Created)
01-AUG-1988 (Rel. 08, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein precursor (Coat polyprotein) (Contains: Knob
GP70: Splke protein P15E).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 7.7%; Score 78.5; DB 1; Length 209; Best Local Similarity 26.5%; Pred. No. 2.3; Matches 50; Conservative 16; Mismatches 56; Indels 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Feline leukemia virus (strain A/Glasgow-1).
Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
NCBI_TaxID=11769;
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        326 IAILGNYSNOTNPPPSCLSTPQHKLTISEVSGGGMCIGTVPKTHQALCNK----TQGGHT 383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        73 GTRIIASPKGPVIQMYINVDKDIJVGWPAPQGSRSLIPPCI------CGSSDI,--- 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----YLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IVLNGAYAQQIRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQIFLATCINGVCWTVYHGA 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels 59; Gaps
                                                                                                                                                                                                              KNOB PROTEIN GP70.

SPIKE PROTEIN P15E.

N-LINKED (GLCNAC. .) (POTENTIAL).

N-LINKED (GLCNAC. .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16; Mismatches 83;
                                                                                                                    EMBL: M12500; AAA43053.1; -.
HSSP; P03385; 1MOF.
InterPro; IPR002050; Env_polyprotein.
Pfam; PF00429; ENV_polyprotein; 1.
SIGNAL

34 445 KNOB PROTEIN GP70.
                                                                                                                                                                                                                                                                                                                                                                                                                                                7.5%; Score 77.5; 1
22.2%; Pred. No. 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search completed: August 30, 2603, 19:13:44
Job time : 10.6577 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GT-GTKALLETAQFROLOMAMHT 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                     71053 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                   445
6425
335
435
91
3307
3307
3307
3307
3307
                                                                                                                                                                                                                                                                                                                                                                                                                  642 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ma.
Local Sim
45;
                                                                                                                                                                                                                                 CHAIN
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13
                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    118
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loc
Matches
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis hepatitis

hepatitis

hepatitis hepatitis hepatitis hepatitis hepatitis

```
SEQUENCE FROM N.A.
MEDLINE-20323484; PubMed-10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
Generation and characterization of a hepatitis C virus NS3 protease-
dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA stage; Flaviviridae;
                        091rs8
091rs3
091rs7
090rs7
090rs7
090rs9
090rs8
091rs9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A Lai V.C.H., Hong 2.;

Lai V.C.H., Hong 2.;

Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.

R HSSP: P6665; JAKP.

HSSP: P6665; JAKP.

R InterPro: JRR000280; CDvir_endptseP80.

R InterPro: JRR001410; DEAD.

R InterPro: JRR001410; DEAD.

R InterPro: JRR001410; HCV_NS3.

R InterPro: JRR0014109; HCV_NS3.

R InterPro: JRR00156; Helicase_C.

R InterPro: JRR001055; WAB_PDNA_binding.

R InterPro: JRR001055; WAB_PDNA_binding.

R InterPro: JRR001055; RNA_pol_DS_PS.

R InterPro: JRR001095; RNA_pol_DS_PS.

R InterPro: JRR001095; RNA_pol_DS_PS.

R InterPro: JRR001094; RNA_pol_DS_VIr.

Pfam: PF002017; HCV_NS3: 1.

Pfam: PF002017; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 4040 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no Pestivirus.
                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Created)
091R83
091R74
091R73
091R73
091R75
091R70
091R70
091R70
091RR9
091RR9
091RR9
091RR0
091RR0
091RR0
091RR0
091RR0
091RR0
091RR0
                                                                                                                                                                                                                                                                                                                      091RS0
091RT8
081754
                                                                                                                                                                                                                                                                                                          Q91RT7
 01-OCT-2000 (TrEMBLrel.
01-OCT-2000 (TrEMBLrel.
01-MAR-2003 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
Genome polyprotein.
Mucosal disease virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11099;
 877
876.5
876
876
Q91FH8
Q91FH8;
                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
Q9IFH8
 DORRER REPRENTED DESTRUCTION OF THE PROPERTY O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nucosal dis
hepatitis c
                                                                                                (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                           ......VAKAVDFIPVESLETTMRSP 195
                                                                                   August 30, 2003, 19:00:22; Search time 37:2105 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091fe5
003463
091rr8
091rr8
091rr3
091rr3
091rr1
091rr6
091rr6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       091fh8
Q81756
                                                                                                                                                                                                                                               830525
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                              Potal number of hits satisfying chosen parameters:
                                                                                                                                                                                                                        830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUMMARIES
                                                                                                                                                                                                                                                                                                      Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                              OM protein – protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091FH8
081756
003465
0091RR8
091RR8
091RR3
091RR3
091RR1
091RR1
091RR6
                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                           1 MKKKGSVVIVGRIVLNGAYA.
                                                                                                                                                                                                                                                                                                                                                                                            sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_mhc:*
sp_organelle:*
                                                                                                                                                                                                                                                                                                                                                                              sp_bacteria:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_phage:*
sp_plant:*
sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                        length: 0
length: 2000300000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_archeap:
                                                                                                                                    US-09-965-594-12
1021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB
                                                                                                                                                                                                                                                                                                                                                                      sp_archea:*
                                                                                                                                                                                                                                                                                                                                                          SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10:11:12:13:14:
                                                                                                                                                                                                                                                                                                           Post-processing:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq
Maximum DB seq
                                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                    Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score
```

..

Database

Result

Searched:

Sequence:

on:

Run

```
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                      NON_TER
SEQUENCE
                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9IFES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
Q9IFE5
    NA WAY TES
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                             10 CINGVCMTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPVGSRSLTPCTGSSDLY; 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CINGVCWTVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYL 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       120 VTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 173
                                                                                                                                                                                                                                                                                                                    29
                                                                                                                                                                                                                                                                                                                                          10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
           SMART: SM00487; DEXDC; 1.
SMART: SM00487; DEXDC; 1.
PROSITE; PS0037; WIRE_1: 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
PROSITE; PS50531; RNASE_T2_2: 1.
ATP-binding: Helicase: Hydrolase; Norstructural protein; Polyprotein; RNA-directed RNA polymerasc: Transferase.
SEQUENCE 4040 AA: 453073 MW; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                                                                    5 GSVVIVGRIVLNG-----AYAQQTRGEEGCQETSQTGRDKNCVEGEVQIVSTAAQTF1,AT
                                                                                                                                                                                                                                                                              5; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                   DB 12; Length 4040;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Choo Q.-L., Richman K., Han J.;
"The nucleotide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL; M32084; AA456777.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                     7;
                                                                                                                                                                                                                               89.0%; Score 908.5; DB 90.8%; Pred. No. 9e-84; Ative 6; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 2436 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD.
InterPro; IPR001531; HCV_NS1.
InterPro; IPR001531; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001869; HCV_NS5.
InterPro; IPR001969; HCV_NSS.
InterPro; IPR0010169; Helicase_C.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PPO1560; HCV_NS1; 1.
Pfam; PPO1578; HCV_NS2; 1.
Pfam; PPO1005; HCV_NS3; 1.
Pfam; PPO1005; HCV_NS4a; 1.
Pfam; PPO1001; HCV_NS4b; 1.
Pfam; PPO1001; HCV_NS5a; 1.
Pfam; PPO1506; HCV_NS5a; 1.
Pfam; PPO1988; Viral_RGRP; 1.
ProDom; PD18605; HCV_NS1; 1.
SWART; SW00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PRINTS; PR00729; CDVENDOPTASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  180 VDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                     Best Local Similarity 90.8 Matches 177; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11103;
                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q81756
Q81756;
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DORENT DE RESTRETAR DE LA COMPANSION DE 
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
5
                                                                                                                                                                                                                                                                                                                                                                              614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                               21
                                                                                                                                                                                                                                                                                           3 KKGSVVIVG---RIVLNG-----AYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                         MEDLINE-21262212; PubMed-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          K
                                                                                                                                                                                                                                                      Gaps
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence..;

Gen. Witzol. 82:1291-1297(2001).

-1. Gen. Witzol. 82:1291-1297(2001).

-1. GOBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                        DB 12; Length 2436;
                                                                                                                                                                                                                                                    Indels
                                                                                                                                       2436 2436
2436 AA; 264734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09IFE5;
0-0-CT-2000 (TrEMBLrel. 15, Created)
01-0-CT-2000 (TrEMBLrel. 15, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                          l.3e-82;
hes 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRI; 3011 AA
                                                                                                                                                                                                                               Pred. No. 1.3e
8; Mismatches
                                                                                                                                                                                                      87.6%; Score 894.5;
85.8%; Pred. No. 1.3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          172 CIRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AF271b34, ....
HSSP; P27958; 1Alv.
InterPro; IPR001410; DEAD.
"APPLO; IPR01410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro: IPR002868; HCV_NSSa.
InterPro: IPR002166; HCV_RdRP.
InterPro: IPR001650; Hellcase_C.
InterPro: IPR007095; RNA_POL_DS_PS.
InterPro: IPR007094; RNA_POL_PSVIF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002531, HCV_NS1.
IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
IPR000745; HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                             Best Local Similarity 85.83
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                       1
2436
```

```
*Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-Jl) with high homology to USA isolates."; Nucleic Acids Res, 20:6410-6410(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87.5%; Score 893.5; DB 1:
85.8%; Pred. No. 2.2e-82;
tive 7; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro: IPR002518; HCV_NS2.
Interpro: IPR004109; HCV_NS3.
Interpro: IPR001409; HCV_NS3.
Interpro: IPR001409; HCV_NS4.
Interpro: IPR001609; HCV_NS5a.
Interpro: IPR002606; HCV_RRP.
Interpro: IPR001650; Helicase_C.
Interpro: IPR001050; Helicase_C.
Interpro: IPR001050; Holicase_C.
Interpro: IPR001050; HCV_RRP.
Pfan: PP01543; HCV_Cappid, I.
Pfan: PP01543; HCV_Cappid, I.
Pfan: PP01543; HCV_Cappid, I.
Pfan: PP01550; HCV_NS2: I.
Pfan: PP01509; HCV_NS2: I.
Pfan: PP01006; HCV_NS3: I.
Pfan: PP01006; HCV_NS3: I.
Pfan: PP01006; HCV_NS3: I.
Pfan: PP010106; HCV_NS4b: I.
Pfan: PP00271; helicase_C: I.
Pfan: PP00271; helicase_C: I.
Pfan: PP0089; VITALIARRP: I.
PFANT: SM0487; DEXDC: II.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
InterPro; IPR002521; HCV_C
InterPro; IPR002519; HCV_E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 85.84
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002531; H
InterPro; IPR002518; H
InterPro; IPR004109; H
InterPro; IPR00145; H
InterPro; IPR00145; H
                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                           STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Inchauspe G.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                              Okamoto
                   SOW WE WANT THE PROPERTY OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1125 CGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21
R Pfam; PF01258; HCV_NS2; 1.
R Pfam; PF01006; HCV_NS3; 1.
R Pfam; PF01006; HCV_NS4s; 1.
R Pfam; PF01006; HCV_NS4s; 1.
R Pfam; PF01001; HCV_NS4s; 1.
R Pfam; PF010998; Viral_RRP; 1.
R ProDom; PD186062; HCV_NS1; 1.
R ProDom; PD186062; HCV_NS1; 1.
R PROSTITE; PS0507; RDRP_POSTITUE; 1.
R PROSTITE; PS50501; RDRP_POSTITUE; 1.
R PROSTITE; PS50521; RDRP_VIRAL; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVG---RIVLNG-----AYAQQTRGEEGCQFTSQTGRDKNQVEGFVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "The 5'-terminal sequence of the hepatitis C virus genome.";
Jpn. J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T., Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87.6%; Score 894.5; DB 12; Length
85.8%; Pred. No. 1.7e-82;
iive 8; Mismatches 10; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane. SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-HC-J1;
MEDLINE-93117120; PubMed-1335573;
Okamoto H., Kanal N., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-HC-J1;
MEDLINE-91013116: PubMed-2170712;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-92044440; PubMed=1658196;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-MAR-2003 (TrEMBLrel. 23, Genome polyprotein. Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity 85.8 les 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            003463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOT ACCOUNT OF THE READ OF THE
               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
1005 RKGREILLGPADGMVSRGWRLLAPITAYAQQIRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      52 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  112 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVG---RIVLNG-----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                   11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
AIP-binding; Coat protein; Envelope protein; Giycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 12; Length 3011;
                                                                                                                                                                                "Characterization and mapping of a B-cell immunogenic domain in
                                                                         STRAIN-HC-J1;
MEDLINE-94174722; PubMed-7510436;
Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase: Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 97E9052C0250463B CRC64;
Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases.
```

```
Hepatitis C
                                                                                                     Matches 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       interPro;
                       SEQUENCE
                                                               Query Match
Best Local 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                 036579
      NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 7
036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       S
                                                                                                                                                                                q
                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ت
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      65 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Genc from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the FMBL/GenBark/DDBJ databases.
BMBL, AF36218; AAR54543.1;
InterPro: IPR004109; HCV_NS3.
Pfam. PF02907; HCV_NS3: i.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18 AYAQQTRGEEGCQETSQTGRUKNQVEGEVQIVSTAAQTFI.ATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN-Pt.1Y:

Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;

"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";

EMBL, SENDITEE (SPR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL, AFS6235; AAR54560.1;

InterPro: IPR040109; HCV NS3.

Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Gaps
                                                                                                                                                                                                                                                                     Viruses; SSRNA positive-strand viruses, no DNA stage: Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssrnA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 891; DB 12; Length 181;
Pred. No. 1.1e-83;
1; Mismatches 5; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 AA
                                                                                                                                             181 AA
                       172 CTRGVAKAVDFIPVESLETIMRSP 195
                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 87.3%;
Best Local Similarity 96.6%;
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Protease.
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091RT5
                                                                                                                                             Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 6
Q91RT5
                                                                                                     RESULT 5
Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TD DE REPERENTED DE LA REPERENTE DE LA REPEREN
                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
0;
                                                                                                                                                                                                                                                 78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                         18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                            5 AYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRII 64
                                                                                                                                                                                                                                                                                                                                            Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone S.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    transcribed RNA.";
Science 277:570-574(1997).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA (BY SIMILARITY).
                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Transmission of hepatitis C by intrahepatic inoculation with
                                                Length 181;
                                                                                                 Indels
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                              87.3%; Score 891; DB 12; 96.6%; Pred. No. 1.1e-83;
                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 | IPR002518; | HCV_NS3.
| IPR0004109; | HCV_NS3.
| IPR000745; | HCV_NS4a.
| IPR001490; | HCV_NS4b.
| IPR001868; | HCV_NS5a.
| IPR002166; | HCV_RRPP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE-97373636; PubMed-9228008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam: PF01001; HCV_NS4b; 1.
Pfam: PF01506; HCV_NS5a; 1.
Pfam: PF00271; helicase_C; 1.
Pfam: PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AF009606; AAB66324.1;
HSSP; P27958; 1HEI.
InterPro; IPR001410; DEAD.
                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein.
                                                                          Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
```

S

```
Q91RR3 PRELIMINARY;
Q91RR3;
Q1-DEC-2001 (TrEMBLrel. 19, Cr
Q1-DEC-2001 (TrEMBLrel. 19, La
Q1-MAR-2003 (TrEMBLrel. 23, La
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              96.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=11103;
                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                 NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                              STRAIN-Pt.4B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                          Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                    Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q91RS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
Q91RS1
  RESULT 9
                      091RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                       ?
                                                                                                                                                                                                                                                            78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           52 AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 AYAQQTRGLLGCIXISLIGRDKNOVEGEVQIVSIAAQTFLATCINGVCWTVYHGAGTRII 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                       3 KKGSVVIVG---RIVLNG----AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gene from
                                                               Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ;
0
                                                                                                                                                          DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP-binding; Coat protein; Envelope protein; Glycoprotein;
Hydrolase; Nonstructural protein; Polyprotein;
RNA directed RNA polymerase; Transferase; Transmembranc.
SEQUENCE 3011 AA; 327182 MW; E2E0EE809c63GLB9 CKC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11arity 96.0%; Pred. No. 1.8e-83; Conservative 2; Mismatches 5;
                                                                                                                                                          87.1%; Score 889.5; DB 1:
84.8%; Pred. No. 5.5e-82;
ive 9; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CTRGVAKAVDFIPVESLETTMRSP 195
                    PROSITE; PS50507; RDRP_POSITIVE; 1. PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                 Matches 173; Conservative
SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-Pt. 3U;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Simines 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
                                                                                                                                                                                  Sest Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
091RR5
10 091RR5
10 091RR5
10 01 -D
10 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                οχ
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
78 ASPRGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             65 ASPKGPVIQMYINVDRDLVGWPAPQGARSLIPCTGGSSDLYLVTRHADVIPVRRRGDSRG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5 AYAQQTRGLLGCIIISLTGRDKNQYEGEVQIVSTAAQTFLATCINGYCHTVYHGAGTRII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                 Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865240; AAR5465.1;
Interpro, IPR004109; HCV-NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         from
                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Holland-Staley C.A., Rovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Ge
Clincal Strains of the Hepatitis C Virus.";
submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865232, AAK54557.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AA; 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 proclease (Fragment).
Hepatitis C virus.
                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 888; DB 12;
Pred. No. 2.3e-83;
2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 888; DB 12;
Pred. No. 2.3e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87.0%; Score 888;
Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87.0%;
96.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity
```

ö

```
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Pt.3T;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                  Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q91RR6
Q91RR6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091RR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
    REPERT NO SECOND SECOND
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
    Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78 ASPKGPVIQMYTNVDKDL,VGWPAPQGSRSLTPCTCGSSDL,YL,VTRHADVIPVRRRGDSRG 137
                                                                                                                                                           65 ASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCTCGSSDLYLVTRHADVIPVRRGDSRG 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 AYAQOTRGLIGCIITSLIGRDKNOVEGEVQIVSTAAQIFLATCINGVCWTVYHGAGIRTI 64
                                             138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                               18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Folland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369245; AAK44570.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
    0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ó.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87.0%; Score 888; DB 12; Length 181; 96.0%; Pred. No. 2.3e-83; Live 2; Mismatches 5; Indels (
    5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 181
181 AA; 19144 MW; COC91F1E2EEB0B32 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                               091ROB;
01-DEC-2001 (TFEMBLrcl. 19, Created)
01-DEC-2001 (TFEMBLrcl. 19, Last sequence update)
01-MAR-2003 (TFEMBLrcl. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AA
                                                                                                                                                                                                                                                                                                                                                                               181 AA
  2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
Matches 170; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBL_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Pt.52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepacivirus
Matches 170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q91RT1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RT1
                                                                                                                                                                                                                                                                                                                                                                             091RQ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
091RT1
091RT
AC 091RT
DT 01-DE
DT 01-DE
DT 01-MS
DE NS3 P
OS HEPAC
OC HEPAC
OX NICBI
RN 13
RN 11
RP, SEQUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                      ò
                                                  õ
                                                                                  셤
                                                                                                                                  ó
                                                                                                                                                                                                                                                                  셤
```

```
ö
                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   78 ASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65 ASPKGPVIQMYINVDQDLVGWPAPQGARSLIPCTGGSSDLYLVTRHADVIPVRRGDSRG 124
                                                                                                                                                                                                                                                                                                                                                                                      77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64
                                                                                                                                                                                                                                                                                                                                                                                                                               5 AYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 AYAQQIRGLLGCIITSLIGRDKNQVEGEIQIVSIAAQIFLAICINGVCWIVYHGAGIRII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    138 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRS 194
                                                                                                                                                                                                                                                                                                                                                                                      18 AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases. EMBL, AF369222; AAR54847.1: -InterPro; IPR004109; HCV_NS3. Pfam. PPC2907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gene from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Ge
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865237; AAK54652.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 181;
                                                                                                                                                                                                                                                                                        Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 886; DB 12; Length 16
Pred. No. 3.7e-83;
3; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 181
181 AA; 19101 MW; 614ADA8BOF33CCAF CRC64;
                                                                                                                                                                                                                    181 181
181 AA: 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                          Score 888; DB 12;
Pred. No. 2.3e-83;
                                                                                                                                                                                                                                                                       87.0%; Scu. 96.0%; Pred. No. 2.. 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86.8%;
                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 96.0
Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
```

```
InterPro;
SO WE WE WANTED THE STREET OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78 ASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCICGSSDLYLVIRHADVIPVRRRGDSRG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The STRAIN-Colone;

STRAIN-Colone;

Desai S.M., Devare S., Yamaguchi J.;

The partitis C Virus.";

"Hepatitis C Virus.";

"Hepatitis C Virus.";

Submitted (Jul-2000) to the EMBL/GenBank/DDBJ databases.

Lish The VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN AND GRYOPO-1;

EMBL: AF290978; AAG02099-1;

RSSP; P27938; HREI.

RSSP; P27938; HREI.

INTERPO: IPRO01410; DEAD.

RICEPPO: IPR0015221; HCV_cose.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                125 SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     138 SLLSPRPISYLKGSSGGPULCPAGHAVGIFRAAVCIRGVAKAVDFIPVESLETTMRS 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18 AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTI
                                                                                                                                                                                                                                                                                                                                                                  STRAIN-PL.174;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369224; AAK5549.1;
InterPro, IPR004109; HCV.NS3.
Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 181
181 AA; 19131 MW; 8BD7FC2769DBD635 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                      01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86.8%; Score 886; DB 12;
96.0%; Pred. No. 3.7e-83;
tive 2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3011 AA
                                                                          181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 96.0
Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                   protease (Fragment).
titis C virus.
                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein. Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                              [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                            Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9ELS8
                                                                       Q91RS9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
Q9ELS8
                       RESULT 14
091RS9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
52 AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KRGSVVIVG --- RIVLNG --- -- AYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE; PS00190: CYTCCHRONE_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Brotein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA: 327107 MW; A6BECF5A3B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 86.7%; Score 885.5; DB 17
Best Local Similarity 84.3%; Pred. No. 1.4e-81;
Matches 172; Conservative 10; Mismatches 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Search completed: August 30, 2003, 19:18:18
Job time : 38.2105 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  172 CTRGVAKAVDFIPVESLETTMRSP 195
                                                Interpro: IPR004109; HCV_NS3.
Interpro: IPR000045; HCV_NS4a.
Interpro: IPR001040; HCV_NS4a.
Interpro: IPR001040; HCV_NS5a.
Interpro: IPR001066; HCV_RSAP.
Interpro: IPR0010650; HCV_RGAP.
Interpro: IPR0010550; Hclicase_C.
Interpro: IPR007095; RNA_pol_DS_PS.
Interpro: IPR007095; RNA_pol_PSvir.
Pfam: PF01543; HCV_Copeid: 1.
                                                                                                                                                                                                                                                                                                        Pfam, PF01539; HCV_ENV; Pfam; PF01539; HCV_ENV; 1. Pfam; PF01539; HCV_ENS; 1. Pfam; PF01539; HCV_ENS; 1. Pfam; PF01001; HCV_ENS4; 1. Pfam; PF01001; HCV_ENS4; 1. Pfam; PF01505; HCV_ENS4; 1. Pfam; PF01505; HCV_ENS5; 1. Pfam; PF00571; helicase_C; 1. Pfam; PF00998; VITAL_RGRP; 1. Pr0Dom; PD186062; HCV_ENS5; 1. SWART; SM00487; DEXDC; 1.
  HCV_NS1.
IPR002531;
IPR002518;
```

Run on:

```
ARI8692 Sequence 10 109329 Sequence 54 109329 Sequence 10 109331 Sequence 10 109331 Sequence 15 M32084 Hepatitis C ARI8728 Sequence ARI18723 Sequence BD09332 HCV culti 10894 Sequence 10894 Sequence ARI76930 Sequence ARI7693 Sequence ARI7648 Sequence BD080334 Hepatitis C ARI8780 Synthetic AF38780 Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AX395309 Sequence
AX54818 Sequence
E08263 gRNA of Hep
E08264 cDNA of Hep
D10749 Hepatitis C
AR127809 Sequence
BD081910 Hepatitis
AR127810 Sequence
BD081911 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR145264 Sequence
AF511948 Hepatitis
AF369218 Hepatitis
AF369235 Hepatitis
                                                                                                                                                                                                                                  AF268278 Pestiviru
                                                                                                                                                                                                                                          AR118686 Sequence
IO6434 Sequence 48
IO9328 Sequence 8
                                                                                                                                                                                                                     AR179057 Sequence
                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ALIGNMENTS
                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                       HPCPLYPRE
AF387805
AF387808
                                                                                                                                                                                                                                                                                          I09329
AR118696
                                                                                                                                                                                                                                                                                                                                        AR118728
AR118722
AR118723
                                                                                                                                                                                                                                                                                                                                                                              108294
AR166930
                                                                                                                                                                                                                                                                                                                                                                                               AR301300
AR176483
BD080334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AF387806
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AF387807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR127809
BD081910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR127810
BD081911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AF369218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AF369235
                                                                                                                                                                                                                                           AR118686
                                                                                                                                                                                                                                                                                                                                                                     BD091382
                                                                                                                                                                                                                                                                                                                                                                                                                                      I71894
I81885
                                                                                                                                                                                                                                                                                                                                                                                                                              E66593
                                                                                            em_htgo_hum: *
em_htgo_mus: *
em_htgo_other: *
                  .htg_inv:*
.htg_other:*
                                             em_htg_pln: *
em_htg_rod: *
em_htg_mam: *
         em_htg_hum:*
                                      em_htg_mus:*
                                                                           em_htg_vrt:*
                                                                                                                                                                                          Query
Match Length DB
                                                                                                                                                                                                                                                                      6785
6785
6785
6785
7310
7310
7310
7310
9105
9185
9185
9185
9185
9185
9185
9401
9401
                                                                                                                                                                                                                                                                                                                                                                                                                                               9401
9401
9609
9609
9618
9646
                                                                                    em_sy:
                           e
E
                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                        (without alignments)
3147.423 Million cell updates/sec
                                                                                                                         1 MKKKGSVVIVGRIVLNGAYA.......VAKAVDFIPVESLETIMRSP 195
                                                                August 30, 2003, 19:18:33 ; Search time 2534.57 Seconds
      GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Etd.
                                              - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                   2888711 scqs, 20454813386 residues
                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                        BLOSUM62
Xgapop 10.0 , Xgapext /
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                      US-09-965-594-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   gb_sts:*
gb_nn:*
gb_un:*
gb_vi:*
em_ba:*
em_fun:*
em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_pat:*
em_ph:*
em_pl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_ro:*
em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_mu:*
em_om:*
em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                    gb_ba:*
gb_htg:*
gb_in:*
gb_om:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  gb_pat:*
gb_ph:*
gb_pl:*
gb_pr:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_in:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        #: vo_dp
                                                                                                                                                                                                                                                                                                                                                                                                                             Gen Emb1:*
                                                                                                                Perfect score:
Sequence:
                                                                                                                                             Scoring table:
                                              OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                             Database :
                                                                                                                                                                                                    Searched:
```

```
Pestivirus
                                                                                                                                                           10864644
                                                                                                                                                                                                                                        source
            VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                              MEDLINE
PUBMED
                                                                                                                                                                    REFERENCE
AUTHORS
                                                                                                                                                                                           TITLE
JOURNAL
                                                                              REFERENCE
                                                                                           AUTHORS
  ACCESSION
                                                                                                                                   JOURNAL
                                                                                                                                                                                                                            FEATURES
                                                                                                               TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VRL 12-JUL-2000
            PAT 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LcuSerProArgProIleScrTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 159
                                                                                                                                                                                                                                                                                                                                        583
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      100 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 179
                                                                                                                                                                                                                                                                                                                                                                         39
                                                                                                                                                                                                                                                                                                                                                                                                                    29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                79
                                                                                                                                                                                                                                                                                                                                                                                                                               20 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGiuThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                       60 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                             GlySerValValIleValGlyArgIleValLeuAsnGly-------AlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                   40 AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr
                                                                                                 Tobases 1 to 12734)
Hong, Z., Lai, V.C.H. and Lau, J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus
Patent: US 6326137-A 1 04-DEC-2001;
Location/Oualifiers
1. 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          180 ValAspPheIleProValGluSerLeuGluThrThrMetArqSer 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
              linear
                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AF268278 12734 bp
Pestivirus type 1, complete genome.
                                                                                                                                                                                2803
       12734 bp
Sequence 1 from patent US 6326137.
AR179057
AR179057.1 G1:20220612
                                                                                                                                                                                                                                                                                                        US-09-965-594-12 (1-195) x AR179057 (1-12734)
                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                              /organism-"unknown"
2604 c 3295 a
                                                                                                                                                                                                                          3.06e-56
928.50
94.87%
92.31%
90.94%
                                                                                         Unclassified
                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                Ø
                                                                              Unknown.
                                                                    Unknown
                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               140
                                                                                                                                                                                                                                                                        Query Match:
                                            VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                          Best Local
                                                                                                             AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                BASE COUNT
                                                                                                                                                                                                                            Pred. No.:
                       DEFINITION
                                  ACCESSION
                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AF268278
                                                                                                                                               FATURES
AR179057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                            DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
Direct Submission
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research
Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LUVUVALLGGRYVLMLLUTYMVLSEOKALĞIOYGSGEVÜMGNLLTHNNIEVUTYFLL
LYLLIREESKKKWULLIYHILVYHPIKSYYTIKLIAGDVYKABSGGGETLGKIDLEFT
LYVLIY 1GLII ARROPTIVYLIX TRAYILILMIGDVYKABSGGGGETLGKIDLEFT
YVVLIY 1GLII ARROPTIVELIY TRAALRYTELTHQEGODI ANANUT TILLMYSYYD
YFRYKKMLQCILSLVSGVFLIRSLIYLGRIEMPEYT IPWRPLTLLLYLLSTITYTR
WRYDYAGLIQCVPILLLYTUMADERFILLILLETYELVKLYTKTYRTRYFDIERSWLGG
IDYTRYDSIYDONGSGEGCVYLFPSROKAGGGRISILLPLIKATLISCVSSKWOLIYMSY
IDYTRYDSIYDONGSEGGVYLFPSROKAGGGRISILLPLIKATLISCVSSKWOLIYMSY
LTLDFMYYMHRKVIEEISGGTNIISRIVAALIELMSWEEEESKGLKKFYLLSGRIRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ASSGRVVGRVKVGKNEESKPTKIMSGIQTVSKNTADLTEMVKKITSMNRGDFKQITLA
TGAGKTTELPKAVIEEIGRHKRVLVLIPLRAAAESVYQYMRLKHPSISFNLRIGDMKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TEGIQRAMFQRGVNRSLHGIWPEKICTGVPSHLATDIELKTIHGMADASEKTNYTGCR
LQRHEWNKHGWCNWYNIEPWILVWNRTQANLTEGQPPRECAVTCRYDRASDLNVVTQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          KNCTPACLPKNTKIVGPGKFDTNAEDGKILHEMGGHLSEVLLLSLVVLSDFAPETASV
MYLILHFSIPQSHVDVMDCDKTQLNLTVELTTADVIPGSVWNLGKWVCIRPNWMPYET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LIIKHKVRNETVASAYGEEEVYGAPKIMTIIKASTLSKSRRCIICTVCEGREWKGGTC
PRCGRHGKPITCGMSLADFEERHYKRIFIREGNFFGMCSRCQGKHRRFEMDREPKSAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PCHISFGSRMPFRQEYNGFVQYTARGQLFLRNLPVLATKVKMLMVGNLGEEIGNLEHL
GWILRGPAVCKKITEHEKCHINILDKLTAFFGIMPRGTTPRAPVRFPTSLLKVRRGLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TGWAYTHQGGISSVDHVTAGKDLLVCDSWGRTRVVCQSNNRLTDBTEYGVKTDSGCPD
GARCYVLNPEAVNISGSKGAVVHLQKTGGEFTCVTASGTPAFFDLKNLKGWSGLPIFE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LPDLDTV IDTGLKCEKRVRVSSK I PFIVTGLKRMAVTVGEQAQRRGRVGRVKPGRYR
SQETATGSKDY HYDLLQAQRYG I EDGI NVTKSFREMNY DMSLY EEDSLLITQLEILNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LLISEDLPAAVKNIMARTDHPEPIQLAYNSY EVQVPVLFPKIRNGEVTDTYENN SENA
ARKLGEDVPVY IYATEDEDLAVDLLGLDWPDPGNQQVVETGKALKQVTGLSSAENALL
                                                                                 Pestivirus type 1
Pestivirus type 1
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVD
QDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRPISYLKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SYSDTKEEGATKKKTQKPDRLERGKMKIVPKESEKDSKTKPPDATIVVEGVKYQVRKK
GKTKSKNTQDGLYHNKNKPQESRKKLEKALLAWAIIAIVLFQVTMGENITQWNLQDNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RDSPTPLTGCKKGKNFSFAGILMRGPCNFEIAASDVLFKEHERISMFQDTTLYLVDGL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TNSLEGARQGTAKLTTWLGKQLGILGKKLENKSKTWFGAYAASPYCDVDRKIGYIWYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GDMATGITYASYGYFCQMPQPKLRAAWVEYSYIFLDEYHCATPEQLAIIGKIHRFSES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /translation="MELNTNEGSGSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YCAECNRLHPAEEGDFWAESSMLGLKITYFALMDGKVYDITEWAGCQRVGISPDTHRV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IRVVAMTATPAGSVTTTGQKHPIEEFIAPEVMKGEDLGSQFLDIAGLKIPVDEMKGNM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LVFVPTRNMAVEVAKKLKAKGYNSGYYYSGEDPANLRVVTSQSPYVIVATNAIESGVT
                                                                                                                                                                                                                           PestLvIrus.
La (bases 1 to 12734)
Lai.V.C., Zhong.W., Skelton,A., Ingravallo,P., Vassilev,V.,
Donis,R.O., Hong.Z. and Lau.J.Y.
Deneration and characterization of a hepatitis C virus NS3
protease-dependent Dovine viral diarrhea virus
J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Pestivirus type
/mol_type="genomic RNA"
/db_xref="taxon:11099"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /codon_start=1
/product="polyprotein"
/protein_id="AAF82566.1"
/db_xref="GI:9049957"
AF268278.1 GI:9049956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2 (bases 1 to 12734)
Lai, V.C.H. and Hong, Z.
```

us-09-965-594-12.rge

```
AR118686.1 GI:14100596
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7.42e-64
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                               .5360
                                                                                                                                                                                                                                                                                               Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                              Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                              1060
                                                                                                                                                                                                                                                                                 Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1167
                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             867
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92
            160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                            RESULT 3
AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                               LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                    Q
O
                                                                           ò
                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                οž
VALEY VU VOALAKRIN PHILTOLYT LEDORE LEDT HILDORAPUNIKTE TOFFIELKELAS

GDVEK INGAL SONA KRAN PHILTOLYT LEDORE AND APPRIATE TOFFIELS I LENKEE

I IRYCLWGTHTALYKS I AARLCHEFAFAT LVIKWIAFGGESVSIJHVKOAAVOLVYYY

MINKPEPPPRESSYVILST TYKYTYE IRKGKESOCIJGTGI SAAMEILSONPOYGI ISW

LGVGAIAAHNA IESSEOKRTLLMKVFVKNFLDOAATDELVKENPEK I IMALFEAVOT

GNPLALTY HLYCYVEWENEARELSEFTAGRIFTETLINEAFEAFELGENOSGGK IRRLEGGY

YILDLYGLIKO INRGLKKWYLGWAA PPESCOMTPSDER I INPALEGOVG

HYYKOYTAK ITDYSKOKMILAPDKWEVENCYTTRLARKTYGOPOMACH LGGOVE

HYYKOYTAK ITDYSKOKMILAPOKNEVENCYTTRLARKTYGOPOMACH LGGOVE

HYYKOYTAK ITDYSKOKMILAPOKNEVENCYTTRLARKTYGOPOMACH LGGOVE

HYYKOYTAK ITDYSKOKMILAPOKNEVENCYTRAN INCOMPERANINE

AYTTVNEDVOTIR KPULGERY IDPOVUTINLOPEVOYDESEVATITIGRETLMTTGOPT

AYTTVNEDVOTIR REVORSTYAMKILSON INCOMPERANINE INCOMPERANINE

VERKTARNINLYTGON PREIR REVOSSTYAMKILSON INCOMPERANINE INCOMPERANINE

VKODWARACHAPTOTOTR I REVOSSTYAMKILSON INCOMPERANINE INTOMPERANINE

VKON MAGANOTOTR I REVOSSTYAMKILSON INCOMPERANINE INTOMPERANINE

YGEVTWEOLDER INRGAAGFLORD INCOMPERANINING INCOMPERANINE

TYCHTY REHWEN INCOMPERANINE INCOMPERANINING INCOMPERANINE

TYCHTY REHWEN INCOMPERANINING OF PRESK INTOMPERANINE

TYCHTY REHWEN INCOMPERANINING OF PRESK INTOMPERANINE

TYCHTY REPROMENTED INCOMPERANINING OF THE REGINAL YNENGEN INTOMPERANINE

TYCHTY REHWEN INCOMPERANINING OF THE REGINAL YNENGEN INTOMPERANINE

TOTH THE TRY WYNEND STANANING OF THE REGINAL YNENGER

THE THE TRY WYNEND STANANING OF THE REGINAL YNENGER

TOTH THE TRY WYNEND STANANING OF THE REGINAL YNENGER

TOTH THE TRY WYNEND OF THE REGINAL YNENGEN INTOMPERANING THE REGINAL YNENGER

TOTH THE TRY WYNEND OF THE REGINAL YNENGEN INTOMPERANING TO THE REGINAL YNENGEN INTOMPERANING TO THE REGINAL YNENGEN INTOMPERANING TO THE REGINAL YNEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuSerProArgProlleSerTyrLeuIysClySerSerGlyGlyProLeuLeuCysPro 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 833 CTTTCGCCCCGGCCCATTTCCTACCTAAAAGGCTCCTCTGGGGGTCCGCTGTTGTGCCCC 892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaGlnGlnThrArgGlyGluGluGlyCysClnGluThrSorG!nThrGlyArgAsplys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----AlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GTTACGAGGCACGCCAACGTCATTCCCGTGCGCCGGCCAGGTGATAGCAGGGGTAGCCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12734
177
6
7
7
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlySerValValIleValGlyArgIleValLeuAsnGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2802 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3293 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-12 (1-195) x AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1.35e-64
908.50
93.85%
90.77%
88.98%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2608 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                           3'UTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ٠.
و
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
1047 GCTGCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGG 1106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GACCAAGACCTTGTGGGCTGGCCGCTCCGCAAGGTAGCCGCTCATTGACACCTGCACT 1226
                                                                                                                                       PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLySAla 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71
             ::::||||
| CGCAGGGGCCGGGAATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---ArglleValLeuAsnGly--
                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                   5360
175
8
10
11
                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                             Mismatches:
Indels:
                                                                                                                                       DNA
                                                                                                                                                                                                                                                                                                                                                ų
                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                           21-NOV-2000;
                                                                                                                                                                                                                                                                                                                                                1145
                                                                                                                                     AR118686 5360 bp
Sequence 53 from patent US 6150087.
AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LysLysGlySerValVallleValGly----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-12 (1-195) x AR118686 (1-5360)
                                                                                                                                                                                                                                                1 (bases 1 to 5360)
Chien,D.Y.
NANNV diagnostics and vaccines
Patent: US 6150087-A 53 21-NOV-
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                 /organism="unknown"
a 1623 c 1532 g
```

	Db 1347 TCGGGGGGTCCGCTGTTGTGCCCCCGGGGCACACACACAC	192	SULT 5 9028 CUS CUS ETMITION Sequence 8 from Pa RESION 109328.1 GI:58796 YWORNS 109328.0 GI:58796 YWORNS 109328.1 GI:58796 YWORNS 109328.1 GI:58796 YWORNS 109328.1 GI:58796 YWORNS 109328.1 GI:58796 YWORNS 1004328.1 io 536 AUTHORS HOUGHCO,M., Choo, JOURNAL Patent: WO 899465 AUTHERS 10094669 AUTHERS 1009469	Alignment Scores: 7.42e-64 Length: Pred. No.: 894 50 Matches Percent Similarity: 89.71% Conscrubert Local Similarity: 87.51% Indels: DB:	US-09-965-594-12 (1-195) x I09328 (1-5360)	Qy 3 LyslysGlySerValVall1eValGly ::::: :::::: bb 867 CGCAGGGGCCGGGAATACTGCTCGGGCCAGC	Oy 18	Qy 32 ThrSerGlnThrGlyArgAspLysAsnGlnVa 	Oy 52 AlaAlaGInThrPheLeuAlaThrCysIleAs	Oy 72 AlaGlyThrArgThrIleAlaSerProLysGl	Qy 92 AspLysAspLeuValGlyTrpProAlaProGl	Oy 112 CysGlySerSerAspLeuTyrLeuValThrar 	Oy 132 ArgGlyAspSerArgGlySerLeuLeuSerPr 	Qy 152 SerGlyGlyProLeuLeuCysProAlaGlyH1
1347 TCGGGGGGTCCGCTGTTGTGCCCCCGGGGGCACGCCGTGGGCATATTTAGGGCCGCGGTG 1406	172 CysThrArgGlyValAlaIysAlaValAspPheIleProValGluSerI.cuGluThrThr 191 	192 MetargSerPro 195 	1 1 4 1 106434	Alignment Scores: 7.42e-64 Length: 5360 Pred. No.: 894.50 Matches: 175 Score: 894.50 Matches: 175 Percent Similarity: 89.78 Mismatches: 10 Query Match: 67.61% Indels: 2 DB:	-965-594-12 (1-195) x 106434 (1-5360)	3 LysLysGlySerValVallleValGly	18	32 ThrSerGinThrGlyArgAspLySAsnGlnValGluGlyGluValGinIleValSerThr 51	52 AlaAlaGinThrPheLcuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71	72 AlaGlyThrArgThrIleAlaSerProLysGlyProvallleGlnMetTyrThrAsnval 91	92 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111	112 CysglySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 131 	132 ArgGlyAspSerArgGlyScrLeuLeuScrProArgProlleSerTyrLeuLysGlyScr 151 	152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGly1lePheArgAlaAlaVal 171
đ	cy Dp	\$ 6	RESULT 4 LOCUS LOCUS DEFINITION VERSION VERSIO	Alignm Pred. Score: Percen Best L Query DB:	-60-SD	oy Op	Qy Db	Oy Ob	Oy Dp	QY Dp	oy op	Qy Dp	O.Y D.b	0y

```
AsnGlyValCysTrpThrValTyrHisGly 71
                                                                                                                                                                                                                                                                                                         heileprovalgluserLeugluthrThr 191
|||||||||||||||||||||||||||||
TIATCCCTGTGGAGAACCTAGAGACAACC 1466
                                                                     PAT 02-DEC-1994
                                                                                                                                                                                                                           HisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                     ----ArgileValLeuAsnGly----- 17
                                                                      linear
                                                                                                                                                                    5360
175
8
10
11
                                                                                                                                                                      h:
es:
rvative:
tches:
                                                                      DNA
                                                                                                                                                3
                                                                                                                        . 6
```

PAT 02-DEC-1994

```
1140 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1199
                                                                                                                                                                                                                                                                                                                                                                                                                            CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArg 131
                                                                                                                                                                                                                                                                                                                                                                                                                   ....------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Thr SerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                               3 LysLysGlySerValVallleValGly-----ArglleValLeuAsnGly---
                                                                                                                           linear
                                                                                                                                                                                                                                                                                                6785
175
8
10
11
                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                                                            DNA
                                                                                                                                                                                                1 (bases 1 to 6785)
Houghton, M., Choo, O.-L. and Kuo, G.
Wanby diagnostics and vaccines
Patent: EP 0318216-A1 54 31-MAY-1989;
                                                                                                                                                                                                                                                                                                  Length:
Matches:
                                                                                                                                                                                                                                                             1429
                                                                                                                         106440 6785 bp
Sequence 54 from Patent EP 0318216.
106440 1 GI:590312
                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-12 (1-195) x IO6440 (1-6785)
                                                                                                                                                                                                                                  Location/Qualifiers
1. 6785
                                                                                                                                                                                                                                                     /organism="unknown"
2050 c 1914 g
                                                                                                                                                                                                                                                                                                9.59e-64
894.50
89.71%
85.78%
87.61%
                                                                                1740 ATGAGGTCCCCG 1751
                                                                      MetArgSerPro 195
                                                                                                                                                                                         Unclassified.
                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                               ď
                                                                                                                                                                       Unknown.
                                                                                                                                                                                 Unknown.
                                                                                                                                                                                                                                                                                           Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     112
                                                                                                                                                                                                                                                                                                                                                                                                                   18
                                                                                                                                                                                                                                                                                                                                                                                                                                                       32
                                                                       192
                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                            ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                              source
                                                                                                                          LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                  TITLE
JOURNAL
                                                                                                                                                                                                                                                              BASE COUNT
                                                                                                                                                                                                                                                                                          Alignment
                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                          AUTHORS
                                                                                                         RESULT 7
I06440
                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                  ď
                                                                     ò
                                                                                      Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCTAACT 1319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 151
                                                                                                                   PAT 16-MAY-2001
                            191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91
                                                                                                                                                                                                                                                                                                                                                                                                                                              21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                            CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                                                                                                                         ---ArgileValLeuAsnGly---
                                                                                                                    linear
                                                                                                                                                                                                                                                                                        6785
175
8
10
11
                                                                                                                                                                                                                                                                                                             Conservative:
                                                                                                                     DNA
                                                                                                                                                                                                                                                                                                                   Mismatches:
Indels:
                                                                                                                                                                                    Chich, D.Y.
Chich, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 65 21-NOV-2000;
Location/Qualifiers
1 .6785
                                                                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                   6785 bp
US 6150087.
                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-12 (1-195) x AR118692 (1-6785)
                                                                                                                                                                                                                                                                                                                                                                        LysLysGlySerValVallleValGly---
                                                                                                                                                                                                                                           /organism-"unknown"
2050 c 1914 g
                                                                                                                   AR118692
Sequence 65 from patent
AR118692
                                                                                                                                            AR118692.1 GI:14100602
                                                                                                                                                                                                                                                                                         9.59c-64
894.50
89.71%
85.78%
87.61%
6
                                                                        11111111111
1467 ATGAGGTCCCCG 1478
                                                               MetArgSerPro 195
                                                                                                                                                                                Unclassified
                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                       Ø
                                                                                                                                                              Unknown.
Unknown.
                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1320
                                                                                                                                                                                                                                                                                                                                                                         m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1560
                                                                                                                                                                                                                                                                                                                                                                                                                                               32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 132
          1347
                                                              192
                            172
                                                                                                RESULT 6
AR118692
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                       source
                                                                                                                                                              SOURCE
                                                                                                                                                                                                                                                                                           ..
O
                                                                                                                                                                                                                                                       COUNT
                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                            PEATURES
                                                                                                                                                                                                          TITLE
                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                       BASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ó
                                                                               g
                                                                                                                                                                                                                                                                                                                                                                        ò
                                            a
                                                              ò
                            ò
```

1619

0y 152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171 b 1620 TGGGGGGTCGCTGTTGCCCCGGGGGCAGCCGTGGGCATATTHGGGCGCGCGGTG 1679 0y 172 CysThrAGGLYVALAlaLySAlaValAspPheIleProValGluSerLeuGluThrThr 191 bb 1680 TGCACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCTGTGGAGAACCTAGAGACAACC 1739 0y 192 MetArgSerPro 195 0y 1940 ATGAGGTCCCCG 1751	RESULT 9 AR118696 LOCUS DECUS DECUS DECUS SET NITION Sequence 74 from patent US 6150087. ACCESSION RELEVANCES VERSION RELYANCES SOURCANISM UNCOWN. ORGANISM ORGANISM ORGANISM ATTHE ATTHE NANBV diagnostics and vaccines JOURNAL Patent: US 6150087-A 74 21-NOV-2006; FEATURES SOURCE TITLE NANBV diagnostics and vaccines JOURNAL Patent: US 6150087-A 74 21-NOV-2006; FEATURES SOURCE 1. 7310 ACCESSION FOR THE NANBV diagnostics and vaccines JOURNAL Patent: US 6150087-A 74 21-NOV-2006; FEATURES SOURCE JOURNAL PATENT US 6150087-A 74 21-NOV-2006; FEATURES SOURCE TO	Alignment Scores: 1.04e-63 Length: 7310 Pred. No.: 894.50 Matches: 175 Score: 894.50 Matches: 175 Percent Similarity: 89.71% Conservative: 9 Best Local Similarity: 85.78% Mismatches: 10 Query Match: 67.61% Gaps: 2	08-09-965-594-12 (1-195) x AR118696 (1-7310) 0y
0y 152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171 111111111111111111111111111111111	RESULT 8	Alignment Scores: 9.59e-64 Length: 6765 Score: 894.50 Matches: 175 Percent Similarity: 89.71% Conservative: 8 Best Local Similarity: 85.78% Mismatches: 10 Ouery Match: 6 Gaps: 2 DB:	

Db 2085 CGGGGTGATAGCAGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCC 2144 Qy 152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyTlePheArgAlaAlaVal 171 Db 2145 TCGGGGGGTCCGCTGTTGTGCCCCGGGGGCACGCCGTGGGCATATTAGGGCCGCGGTG 2204	Oy 172 CysThrArgGlyValAlaLysAlaValAspPheileProValGluSerLeuGluThrThr 191	Qy 192 MetargSerPro 195	DECEMBER 1000 PARTIES OF VITUS POLYPTOTEIN GENE, PARTIES CIGS. MACCESSON MAJOR. 2 1329875 MACCESSON MAJOR. 3 1329875 MATHER PREMEMER 100 PARTIES CHING MANUALIS CONTROLLING. CHINGAL AND MALOUTED PROJECT CHINGAL SOUNCE MAJORAL SOUNCE	THOTHE ADOMEST WILL MORE THOU THUT MODETT OU ANNUT HUT AGE UT CUMBLUGGTE
Db 2085 CGGGGTGATAGCAGGGCCAGCTGCTGTCGCCCGGCCCATTTCCTACTTGAAAGGCTCC 2144 Qy 152 SerGlyGlyProLeut.cuCysProAlaGlyHisAlaValGlyIlePhcArgAlaAlaVal 171 LITTLITTLITTLITTLITTLITTLITTLITTLITTLI	Qy 172 CysThrArgGlyValAlaLySAlaValAspPheIleProValGluSerLeuGluThrThr 191 	Oy 192 MetargSerPro 195 	103131 1	

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              152
                                                                                                                                                                                                                                                                                                                                             ~
                                                                                                                                                                                                                                                                                                                                                                                                                                          32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            192
                                                                                                                                                                                                                                                                                                                                                                                           18
                                                                                                                                                                   source
AR118703
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                     Pred. No.:
                                                                                 ORGANISM
                                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                        FEATURES
                                                                       SOURCE
                                                                                                                                                                                                                                                   Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
         PEFFELDGURLHRPAPPCRPLIREEVSFRUGLHEYPVGSQLPCEPEPDVAVLTSMLT
DPSHITABAAGRIAAGSPPSVASSSASQLSAPSLAATCTAHHDSPDAELIBANLJWR
QEWGGNITRVESENRVYILDSFDPLVAEEDEREISVPAETLRKSRRFAQALPVWARPD
YNPPLVETWKRPDVEDPPWROCPLPPPKSPPVPPPRKRTVVJTGSTLSTALAELATR
SPSSSTGSITGDNTTTSSEPAPSGCPPDSDAESYSMPPLGGEDPDLSOSWSTV
SSENDSTGSTGDNTTTSSEPAPSGCPPDSDAESYSMPPLGGEDPDLSOSWSTV
SSENDSCYGLOSHYOVYLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGTG
QRQKKVTPORLQVLOSHYOVYLKEVKAAASKVKANLLSVEEACSLTPPHSAKSKFGTG
AKDVRCHARKAVTHINSVWKDLLEDNYTPIDTIMAKNEVFCVOPEKGGRRPRLIVF
PDLGVYRCEKMALTVOVTKLPLENAMGSYGFQYSFGORVEFLVOAWKSKTPMGFSYD
TRCPDSTYTESDIRTEEATYCOCDLDPOARVAIKSLTEKLYVGGPLTMSGENGGYRR
CRASGVLTTSGGNTLTCYIKARAACRAGLOCTMLVCGDDLVVICESAGVQEDAASL
RAFTERATRYSAPPODPQPPISDLLITSCSSNVSVAHDGAGKRYYTLTRDPTPLAR
AARTARHTPVNSULGNIIMFADTLWARMILMTHFFSVLIARDQLEGIYGACY
SIEPLDLPPIIQRL
STEDLDLPPIIQRL
STEDLDLPPIIQRL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACT 1844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
  PINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTTDNLKCPCQVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         92 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallLeProValArgArg 131
                                                                                                                                                                                                                                                                                                                                                                                                                              21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7.1
                                                                                                                                                                                                                                                                                                                                                                                ......AlaTyrAlaGlnGlnThrArgGlyGluGiuGlyCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ArgGlyAspSerArgGlySerLeuLeuSerProArgPro1leSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2085 CGGGGTGATAGCAGGGGCAGCCTGTTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2205 TGCACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACCACC
                                                                                                                                                                                                                                                                                                                                   LysLysClyServalVallleValGly-----ArglleValLeuAsnGly----
                                                                                                                                                                                                                                                                                                                                                                                                                              ThrSerGlnThrGlyArgAspLySAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                          7310
175
8
10
11
                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                              2058 9 1539
                                                                                                                                                                                                                                                                                                            US-09-965-594-12 (1-195) x HPCPOLYP (1-7310)
                                                                                                                                                                                                                         1.04e-63
894.50
89.71%
85.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2265 ATGAGGTCCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MetArgSerPro 195
                                                                                                                                                                             2218 c
                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                1495
                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                  ٣
                                                                                                                                                                                                                                                                                                                                                                                                                               32
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                172
                                                                                                                                                                                                                                                                                                                                                                                 18
                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ογ
```

```
2730
                                                                                                                                                                                                                                                                                                                                                                                         2790
                                                                                                                                                                                                                                                                                                                                                                                                                                    2850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2970
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   191
                                                                                                                                                                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                                                                          77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91
                                                                                                                                                                                                                                                                                                                                                                   31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2671 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                               SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    172 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                           AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                   -------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                       ---ArgileValLeuAsnGly
     linear
                                                                                                                                                                                                                   8316
175
8
8
10
11
                                                                                                                                                                                                                               Matches:
Conservative:
Mismatches:
Indels:
                                                                                          (bases 1 to 8316)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 88 21-NOV-2000;
Location/Qualifiers
8316 bp 1
Sequence 88 from patent US 6150087.
                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                       LysLysGlySerValValIleValGly----
                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                 US-09-965-594-12 (1-195) x AR118703 (1-8316)
                                                                                                                                                              /organism="unknown"
2529 c 2345 g
                                      AR118703.1 GI:14100613
                                                                                                                                                                                                                    1.2e-63
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MetArqSerPro 195
                                                                                                                                                   .8316
                                                                                   Unclassified.
                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                         1671 a
                                                       .
Unknown.
                                                                         Unknown
```

σ

```
3811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3871
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3931
                                                                                                                                                                                                                                                                                                                                                                                   3571
                                                                                                                                                                                                                                                                                                                                                                                                                    3631
                                                                                                                                                                                                                                                                                                                                                                                                                                                       3691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3751
                                                                                                                                                                                                                                                                              332 CGCAGGGGCCGGGAGATAÇTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGTGGAGG 3391
                                                                                                                                                                                                                                                                                                                3451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           171
                                                                                                                                                                                                                                                                                                                                                                                                                                     111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       131
                 PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                  7.1
                                                                                                                                                                                                                                                                                                                                                                                                    91
                                                                                                                                                                                                                                                                                              31
                                                                                                                                                                                                                                                             17
                                                                                                                                                                                                                                                                                              52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                            CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                            LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly-----
                 linear
                                                                                                                                                                                 9185
175
8
10
11
                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                  1 (bases 1 to 9185)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 122 21-NOV-2000;
Location/Qualifiers
1. 9185
                9185 bp | from patent US 6150087.
                                                                                                                                                                                                                                           (1-9185)
                                                                                                                                       /organism="unknown"
2790 c 2608 g
                                                                                                                                                                                                                                            US-09-965-594-12 (1-195) x AR118722
                       AR118722
AR118722
AR118722.1 GI:14100632
                                                                                                                                                                                  1.33e-63
894.50
89.71%
85.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MetArgSerPro 195
                                                                             Unclassified.
                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                 ಳ
                  AR118722
                                                                                                                                                1849
                                                            Unknown
                                                                      Unknown
                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 192
                                                                                                                                                                                                                     Ouery Match:
RESULT 14
AR118722
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                    Pred. No.:
                                                                                   REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                       FEATURES
                                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                             ď
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 191
                                                                                                                                                                                                                                                                                                                                                                                                                              AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgScrLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 131
          PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                           7.1
                                                                                                                                                                                                                                                                                      31
                                                                                                                                                                                                                                                                                                                                                           52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                        -------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                 LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly--
          linear
                                                                                                                                                                          8987
175
8
10
11
                                                                                                                                                                                           Conservative:
Mismatches:
Indels:
                                                                  Unclassified.
1 (bases 1 to 8987)
2 (bien, D.Y.
NAMBV diagnostics and vaccines
Patent: US 6150087-A 137 21-NOV-2000;
Location/Qualifiers
                                                                                                                                                                           Length:
Matches:
          AR118728 8987 bp 5
Sequence 137 from patent US 6153087.
AR118728.1 GI:14100638
                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                      US-09-965-594-12 (1-195) x AR118728 (1-8987)
                                                                                                                               /organism~"unknown"
2735 c 2547 g
                                                                                                                                                                          1.3e-63
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MetArgSerPro 195
                                                                                                                          .8987
                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                           a
                                                    Unknown.
                                                              Unknown
                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3433
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         192
                                                                                                                                                                                                                                                                                                                                                                                                                                               3313
                                                                                                                                                                                                                                                                                                                                          3133
                                                                                                                                                                                                                                                                                                                                                                             3193
                                                                                                                                                                                                                                                                                                                                                                                              72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 112
                                                                                                                                                                                                                                                       m
                                                                                                                                                                                                                                                                                        18
                                                                                                                                                                                                                                                                                                                          32
                                                                                                                                                                                                                                                                                                                                                                                                              3253
                                                                                                                                                                                                                                                                                                                                                                                                                                92
                                                                                                                                                                                                            Query Match:
DB:
                          ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                          source
          LOCUS
DEFINITION
                                                                                                                                                                             ..
9
                                                                                                                                          BASE COUNT
                                                                                     AUTHORS
TITLE
JOURNAL
                                                                              REFERENCE
 AR118728
                                                                                                                FEATURES
                                                                                                                                                  ORIGIN
                                                                                                                                                                             Pred.
```

Q ò g ò QQ ò g ò qq ŏ g ò qq ò

8

ò

qq

ò

oy Db ò

```
Search completed: August 31, 2003, 00:45:40 Job time : 2558.57 secs
            음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5374 TCGGGGGGTCCGCTGTTGTGCCCCGCGGGCACGCCGTGGGCATATTTAGGGCCGCGGTG 5315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5794 TIGCIGGCGCCCATCACGCGTACGCCCAGCAGAACAAGGGCCTCCTAGGGTGCATAATC 5735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5494 TGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCGCGG 5435
                                                                                                                                                                                                                                                                                                                                                                                          5854 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 5795
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                112 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             171
                                              PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          172 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AlaGlyThrArgThr11eAlaSerProLysGlyProVal11eGlnMctTyrThrAsnVal 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                             32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 51
                                                                                                                                                                                                                                                                                                                                                                      3 LysLysGlySerValValIleValGly----- ArgIleValLeuAsnGly----- 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                              linear
                                                                                                                                                                                                                                                              9185
175
10
10
11
                                                                                                                                                                                                                                                              Length:
Matches:
Conscrvative:
Mismatches:
Indels:
                                                                                                                          Unclassified.

1 (bases 1 to 9185)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 123 21-NOV-2000;
1. 9185
                                            AR118723 9185 bp D
Sequence 123 from patent US 6150087.
AR118723.1 G1:14100633
                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                               US-09-965-594-12 (1-195) x AR118723 (1-9185)
                                                                                                                                                                                                       /organism="unknown"
a 2608 c 2790 g
                                                                                                                                                                                                                                                               1.33e-63
894.50
89.71%
85.78%
87.61%
3932 ATGAGGTCCCCG 3943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MetArgSerPro 195
                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                Unknown.
                                                                                                                Unknown
                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        192
                                                                                                                                                                                                                                                                                                                                                                                                                  18
                                                                                                                                                                                                source
                                                       DEFINITION
ACCESSION
VERSION
KEYWORDS
                      RESULT 15
AR118723/c
LOCUS
                                                                                                                ORGANISM
                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                  ..
УО
                                                                                                                                     REFERENCE
AUTHORS
TILE
                                                                                                                                                                       JOURNAL
                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Οŷ
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

(without alignments)
2905.924 Million cell updates/sec MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETTMRSP 195 ; Search time 181.082 Seconds protein - nucleic search, using frame_plus_p2n model August 30, 2003, 19:13:57 US-09-965-594-12 Title: Perfect score: Sequence: Scoring table: Run on:

0.5 7.0 7.0 BLOSUM62 Xgapop 10.0, Xgapext Ygapop 10.0, Ygapext Fgapop 6.0, Fgapext Delop 6.0, Delext

2552756 seqs, 1349719017 residues Searched:

Total number of hits satisfying chosen parameters:

5105512

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Command line parameters:

-WODEL-Itame+_pro.model -DEV-xlp
'Q-Cap21_VICSPTO_SPOOL/VE09965594/runat_29082003_151918_28302/app_query.fasta_1.2872
'Q-Cap21_VICSPTO_SPOOL/VE09965594/runat_29082003_151918_28302/app_query.fasta_1.2872
-QF072_IVOSPTO_SPOOL/VE09965594/runat_app_soff="runat" - NaTHIX-blosum62 -TRANS-humana40.cdi
-LIGAT-45 -DGCALIGH-200 -THR_SCORE-PROCT -THR_MRAX-100 -THR_MRAD-O -LIGGN-15
-USER-USO9965594_eCGN_1_1-1412_@runat_29082003_151918_28302 -NCPD-6 -ICPD-3
-NO_MARAP -LARGEQUERY -NEGLSCORES-0 -WAXI -DSPSELCK*100 -LONGLOG
-FGAPOP-120 -WAXI -ILA12_@runat_29082003_151918_28302 -NCPD-6 -ICPD-3
-NO_MARAP -LARGEQUERY -NEGLSCORES-0 -WAXI -DSPSELCK*100 -LONGLOG
-FGAPOP-120 -WAXI -ILAHDOUT-30 -THREADS-1 -XAAPOP-10 -XGAPEXT=0.5 -FGAPOP-6
-FGAPEXT=7 -YGAPOP-10 -YGAPEXT=0.5 -DELEXT-7

Database

**Sides January | Sides | Side /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:* .DS1/gcqdata/qeneseq/geneseqn-emb1/NN1999.DAT:* DS1/gcgdata/genescq/geneseqn-emb1/NA2000.DAT:* /SIDS1/gcgdata/geneseq/genesegn-emb1/NA2001A.DAT: /SIDS1/gcgdata/geneseg/genesegn-embl/NA2003.DAT:* N_Geneseq_19Jun03: 14: 15: 16: 17: 19: 20:

No. is the number of results predicted by chance to have a greater than or equal to the score of the result being printed, score Pred.

is derived by analysis of the total score distribution.

SUMMARIES

		æ			SOUTHWEE	
Result No.	Score	tcl	Length	DB		escription
	102	: 0	885			
2	O	. ~	0	12	333	paritie o viru
m	8	ω.	۱ œ	21	332	Litis C viru
4	œ	9	6	21	333	itis C viru
S	Ω	ω.	9	21	7333	titis C viru
9	3	Э.	9	21	333	citis C viru
_	S.	ς.	9	21	7333	citis C viru
ω.	85	4	59	21	7333	r is
6	28.		3	24	561	eric BVDV/
10	18		61	25	570	Anti-viral synthet
11.		۲.	0	01	an a	ined open
77	4.	٠.	36	2 :	9032	is C vir
21.		٠.	8	01	9210	ed open r
.	, ,	٠.	7	01	7710	ed oben rea
17	4.	٠.	ς;	9	<u> </u>	Sit
91.	4.	· .	ς,	16	2	itis C vi
17	4.	٠.	3	21	<u>ج</u>	sednence
18	4.	٠.	13	50	5	ectide sequ
19	94	٠.	18	11	55	strand
20	94	۲.	18	12	29	titis C vir
77		٠,	8	21	გ:	strand
77	4.	٠,	9	13	2174	lled HCV CDN
23.0	4.	٠.	9	17	1271	atitis C viru
47		٠,	9	9,0	966	polyprotein
67		∵.	2 :	7 0	9	yprotein
970	ر د د	٠,	2 5	77	Ď,	itis C v
77	D	٠,	n d	77	400	atitis C
9 0	مر	٠,	ט ל	7 (NS3/4a
67	λ (٠,	Si	4 4	4	C viru
9.5	מ מ	٠,	S	7,	ABX14410	oding HCV
7 6	٠.	٠,	2 5	27	` '	atitis C vir
3.5	ر د	٠,	6	0.70	325	DNA.
3.5	8		7	20	25	smid pET
34			65.	50	32	NS4A-NS3
3	7.		₿,	70	673	ģ
36	91.	7	18	20	0045	itis C
37	90.	ζ.	31	11	595	to
38	9	۲.	90	24	3450	itis C v1
38	83	ζ.	90	24	3176	itis C viru
0.4	89.	۲.	64	19	2836	atitis C vi
41	89.	۲.	964	24	8728	encoding he
42	89.	7	298	19	5936	tis C vi
43	889.5	87.1	12980	24	ABK87286	titis C viru
44	89.	۲.	662	21	3621	lectide sequen
4.5	88.	7.	99	20	8035	NS4A-NS3

ALIGNMENTS

Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds. Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2. BP AAA73329 standard; DNA; 588 (first entry) 19-DEC-2000 AAA73329; RESULT 1 AAA73329

Hepatitis C virus. Synthetic.

Location/Qualifiers

Key

N

```
CCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTACCTGGTT 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              P-PSDB; AAB15221.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
                                                                                                                                               161
                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                       AAA73330
                                                                                                                                                                                                                                                              RESULT 2
                                                           D
                                                                                       ò
                                                                                                                   a
                                                                                                                                                 ò
                                                                                                                                                                           qq
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                 QC.
                                 õ
                                                                                                                                                                                                                                                                                                                                                                            The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVal 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 CAGCAGACTCGAGGTGAGGAGGGTTGCCAAGAACCTCCCAGACCGGTCGTGACAAAAAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LysGlyProValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAla 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                           Modified hepatitis C virus (HCV) NS3 protease comprising at least l substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis \mathbb C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GinginThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLysAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41 GlnValGluGlyGluValGlnIleValSerThrAlaAlaGinThrPheLeuAlaThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 IleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro
1..588
/*tag= a
/product* "NS4A-NS3 fusion protein #2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 588
0
0
0
0
0
0
                                                                                                                                                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-12 (1-195) x AAA73329 (1-588)
                                                                                                                                                                                                       Shang Y,
                                                                                                                                                                              ဥ
                                                                                                                                                                                                                                                                                                                                                      Claim 26; Fig 12; 66pp; English
                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4.5e-84
1021.00
100.00%
100.00%
                                                                                                                                                 99US-0115271
                                                                                                                                                                                                       Wittekind M, Weinheimer S,
                                                                                                                    2000WO-US00345
                                                                                                                                                                                                                                    WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                  P-PSDB; AAB15220
                                                            WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
Pred. No.:
                                                                                                                                                 08-JAN-1999;
                                                                                                                    06-JAN-2000;
                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                      SerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAla 160
                                                                                                                                                   GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaVal 180
                                                                                                                                                                                        481 GGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAAGCTGTT 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                  ThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLeu
                                                                                                               421 TCCCCGCGTCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGCTGTGCCCGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3.
                                                                                                                                                                                                                                                AspPhelleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /*tag= a
/product= "NS4A-NS3 fusion protein #3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
1..594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 26; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                               ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Weinheimer S,
                                                                                                                                                                                                                                                                                                                                               AAA73330 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000 (first entry)
```

us-09-965-594-12.rng

```
ThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLeu 140
                                                                                      P-PSDB; AAB15212
    WO200040707-A1
                                         08-JAN-1999;
                                                                                                                                                                                                                                    Alignment Scores:
                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                  41
                                                                                                                                                                                                                                                                                                                                                                                                    81
                                                                                                                                                                                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                         301
                                                                                                                                                                                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                          ..
Q
                                                                                                                                                                                                                                          Pred. N
Score:
à
                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                          118
                                                                                                                                             180
                                                                                                                                                                     240
                                                                                                                                                                                              300
                                                                                                                                                                                                                       360
                                                                                                                                                                                                                                    138
                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                             158
                                                                                                                                                                                                                                                                         480
                                                                                                                                                                                                                                                                                                  540
                                                                                                               ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 178
                                                                               18
                                                                                           9
                                                                                                       38
                                                                                                                                 58
                                                                                                                                                         78
                                                                                                                                                                                  8
                                                                                                                                                         CGGCTCCGCAGGGTTCCCGTGCCTGACCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                    TyralaGlnGlnThrargGlyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                          ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                             LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                               MetLysLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly-----Ala
                                                                                                                                       protease; viral replication; chronic liver disease;
liver cancer; ds.
                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1
    T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             /product- "NS3-NS4A fusion protein"
                      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
    G; 149
                                                                   (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
    C; 156
                                                                   US-09-965-594-12 (1-195) x AAA73330
                                                                                                                                                                                                                                                                                                                                                  ВР
                     5.64e-82
998.00
98.48%
97.97%
    BP; 103 A; 186
                                                                                                                                                                                                                                                                                                                                                  AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                        /*tag-
                                Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                           liver failure;
   Sequence 594
                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000
                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                        13
                                                                                                                    19
                                                                                                                                 39
                                                                                                                                             121
                                                                                                                                                                     181
                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                    119
                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                            139
                                                                                                                                                                                                                                                                         421
                                                                                                                                                                                                                                                                                      159
                                                                                                                                                                                                                                                                                                  481
                                                                                                                                                         59
                                                                                                                                                                                                          66
                                                                                                                                                                                                                                                                                                                                                              AAA73328;
                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                           Key
                                                                                                                                                                                                                                                                                                                                            AAA73328
                                                                                                                                                                                                                                                                                                                                      RESULT
                              Score:
                                                                                                                                                                                                                                                                                                                                                   g
                                                                               ö
                                                                                           g
                                                                                                        ò
                                                                                                                    셤
                                                                                                                               ò
                                                                                                                                            셤
                                                                                                                                                         ŏ
                                                                                                                                                                  셤
                                                                                                                                                                                 ö
                                                                                                                                                                                             g
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                    요
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                          ò
```

```
The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAla 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 GlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLysAsn 40
                                                                                                                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IleasnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         588
000
000
000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                        >
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                        Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-12 (1-195) x AAA73328 (1-588)
                                                                                                                                      Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                     Disclosure; Fig 10; 66pp; English.
                                                                                           ဗ
                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1.59e-80
982.00
97.44%
97.44%
                                           99US-0115271.
D6-JAN-2000; 2000MO-US00345
                                                                                                                                                                                   WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
```

240

80

300

120

9

9

31 15:59:03 2003

Sun Aug

1.99e-80

```
..
OM
                  Score:
    Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                        ç
                                                                                                                                                  g
                                                                                                                                                                                                  a
                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mutants and NS3-NS4A fusion proteins which can be used to identify thibbliors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-halixo-1 variant.
                                                                                            The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3
361 ACCCGTCACGCTGACGTTATCCCGGTTCGTCGTGGTGACTCCCGTGGTTCCCTGCTG 420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                          GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLySAlaVal
                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                    181 AspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                  /product= "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
1..594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M, Weinheimer S,
                                                                                                                                                                                                                 AAA73331 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-JAN-2000; 2000WO-GS00345
                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
P-PSDB; AAB15222.
                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                            AAA73331;
                                                    421
                                                                              161
                                                                                                         481
                                                                                                                                                                                                  AAA73331
ID AAA
                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                               ö
                                                    a
                                                                              õ
                                                                                                      a
```

Sequence 594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;

Aliqnment Scores:

```
99 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 118
                                                                                                                                                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                301 CAGGCTCCGCAGGGTTCCCTGTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                   LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 158
                                                                                                                1 MetLysLysLysGlySerValVallleValGlyArglleValLeuAsnGly-----Ala 18
                                                                                                                                                                                                                                                                                                                   SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 98
                                                                                                                                       1 ATGAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT 60
                                                                                                                                                                                                                                                                  59 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla 78
                                                                                                                                                                                                                   LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                 protease; viral replication; chronic liver disease; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /*tag= a
/product= "NS4A-NS3 fusion protein #8"
   594
190
                          Conservative:
Mismatches:
Indels:
Length:
Matches:
                                                                                       US-09-965-594-12 (1-195) x AAA73331 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA73335 standard; DNA; 594
              981.00
96.95%
96.45%
96.08%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
Synthetic.
                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3
liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                     119
                                                                                                                                                                                                                 39
                                                                                                                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA73335;
                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA7333
```

```
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                               WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 594
                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
                                                      481
                                159
                                                                                                                                          ò
                                                      g
                                                                         ò
                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                118
                                                                                                                                                                                                               fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV gonome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CCGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                       sequence is the coding sequence for a mutated version of a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TyralaGlnGlnThrargClyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ThrCys11eAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MetLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly-----Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sequence contains the alpha-helix0 wild-type sequence
                                                                                                                                                                                                                                                                                                                                                                               594
1188
12
2
                                                                                                                                                                                                                                                                                                                                               Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                          Wittekind M, Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-12 (1-195) x AAA73335 (1-594)
                                                                                                                                                                                Disclosure; Fig 18; 65pp; English.
                                                      (BRIM ) BRISTOL-MYERS SQUIBB CO
           2000WO-US00345.
                                99US-0115271
                                                                                                                                                                                                                                                                                                                                                                             2e-78
959.00
95.94%
95.43%
                                                                                               WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                          P-PSDB; AAB15226
           06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                    Aliqnment Scores:
                                08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
```

```
mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 178
                                                                                                                                                          CCGCCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 fusion protein created using the Hepatitis C virus (HCV) N33 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The present sequence is the coding sequence for a mutated version of a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                  179 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5.
                                                                                                                                                                                                                                                      other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                594
187
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BP; 105 A; 189 C; 153 G; 147 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protease and protease:inhibitor complexes. The sequence contains the alpha-helix0-1 variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 26; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ပ္ပ
                                                                                                                                                                                                                                                                                                                                                                                                             ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.07e-77
951.00
95.43%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                             AAA73332 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                P-PSDB; AAB15223.
```

```
240
                                                                                                                                                                180
                                                                                                                                                                                                                                                         LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 158
                                                                                                                                                                                                                                                                                      ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 178
                                                                          38
                                                                                                       28
                                                                                                                                    78
                                                           9
                                             TyrAlaglnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                       LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                     ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                           LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
1..594
/*teg a
/product = "NS4A-NS3 fusion protein #6"
727
Mismatches:
       Indels:
Gaps:
                              US-09-965-594-12 (1-195) x AAA73332 (1-594)
                                                                                                                                                                                                                                                                                                                                                                BP
                                                                                                                                                                                                                                                                                                                                                                AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
94.928
93.148
21
                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
Synthetic.
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-JUL-2000
                                                                                                                     121
                                                                                                                                    29
                                                                                                                                                                                                                                          361
                                                                                                                                                                                                                                                                                                    481
                                                                          19
                                                                                                       39
                                                                                                                                                                                                                           119
                                                                                                                                                                                                                                                         139
                                                                                                                                                                                                                                                                       421
                                                                                                                                                                                                                                                                                      159
                                                                                                                                                                                                                                                                                                                                                                              AAA73333;
                                                                                        61
         Query Match:
                                                                                                                                                                                                                                                                                                                                                 RESULT 7
                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                   윱
                                                                                                                                                q
                                                                                                                                                                 ò
                                                                                                                                                                               g
                                                                                                                                                                                             ò
                                                                                                                                                                                                          q
                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                  g
                                             ò
                                                          g
                                                                          à
                                                                                                    õ
                                                                                                                                   ò
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     101 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MetLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly-----Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          39 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             594
1184
2
2
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
                                                             Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-12 (1-195) x AAA73333 (1-594)
                                                             Υ,
                                                             2hang
                                ဥ
                                                                                                                                                                                                                 Claim 26; Fig 16; 66pp; English
                                (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.32e-76
939.00
94.92%
93.40%
91.97%
 99US-0115271
                                                             Weinheimer S,
                                                                                          WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                             Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      용
 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
```

28

18 9 38 240

86

78

360

420

119 361 139 421

ò 셤 ò

셤

~

AAA7333 RESULT

```
19 TyralaGlnGlnThrArgGlyGluGluGlyGySGlnGluThrSerGlnThrGlyArgAsp 38
                                                                   LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                           SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                             421 CIGCIGICCCCGCGICCCGAICICCIGAAAGGIICCICCGGIGGICCGCIGCIGIGC
                                                                                                                                                                                      39 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                          ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                        AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MetLysLysLysGlySerValValIleValGlyArgIleValLeuAsnGly--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pestivirus; Npro; protease; NS3; screening; ds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Chimeric – Bovine viral diarrhea virus.
Chimeric – Hepatitis C virus.
               JS-09-965-594-12 (1-195) x AAA73334 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04-DEC-2001,
                                                                                                                                                                                                                                                        59
                                                                                                                                                                                                                                                                                                                           19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABA95615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hong Z,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABA9561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                   g
                                                                                                                                                                                        ò
                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                 g
                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both wessential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLys 178
                               CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGGTTGCTAAA 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                          Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                   179 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 195
                                                                                     Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     594
183
3
9
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
1..594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 26; Fig 17; 66pp; English.
                                                                                                                                                                     BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.08e-75
929.00
94.42%
92.89%
90.99%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                     AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0115271
                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-465976/40.
P-PSDB; AAB15225.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38-JAN-1999;
                                                                                                                                                                                                                                        19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                               Synthetic
                                                                                                                                                                                                      AAA73334;
                                 481
```

178

Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)

WPI; 2002-121103/16.

158 480

300

86

78

us-09-965-594-12.rng

(first entry)

28-MAR-2003

ABX15706;

```
pestivirus genome where the Npro protease genc is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
```

Example 2; Columns 17-28; 20pp; English.

The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct encomprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a genome encoding a functional HCV NS3 protease. Furthermore, lunction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

```
Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                      US-09-965-594-12 (1-195) x ABA95615 (1-12734)
     5.05e-74
928.50
94.87%
92.31%
                                   Similarity:
                         Percent Similarity:
Alignment Scores:
                                                                                       S.
                                           Query Match:
                                   Best Local
                                                                                                        g
                                                                                                                                           셤
                                                                                        ŝ
                                                                                                                                                              ö
                                                                                                                                                                               셤
```

LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 159 113 GGTAGTGTTGTTATTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 179 79 66 39 59 ProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro CTTTCGCCCCGGCCCATTTCCTACCTAAAAGCTCCTCGGGGGGTCCGCTGTTGTGCCCC 60 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer GlyServalvalilevalGlyArgilevalLeuAsnGly-------AlaTyr 20 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 40 AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 194 80 553 100 713 773 140 833 893 180 953 120 a 셤 à 셤 ద ö g g a ö ð à ò õ

ABX15706 standard; DNA; 612 BP. RESULT 10 ABX15706

612 179 5 6 5

Length: Matches: Conservative: Mismatches:

1.01e-74 918.50 94.36% 91.79%

Best Local Similarity:

Query Match:

Percent Similarity:

Alignment Scores:

Indels:

Gaps:

US-09-965-594-12 (1-195) x ABX15706 (1-612)

```
This invention relates to a novel synthetic viral prototoxophore domains a toxin moiety operatively incorporated into a substrate domain specific for a viral enzyme. This prototoxophore may be bound and moistled by the viral enzyme thus converting it to a toxophore. Also disclosed in the invention is a method for enhancing the anti-viral enter this method comprises contacting a cell, effected with a virus or is susceptible to infection, with a prototoxophore. The invention further comprises an assay to identify attribute agents, comprising contacting an infected cell with a candidate agent and comparing the ability of the agent to inhibit the growth or infectivity of the virus in the cell. The prototoxophores of the invention may be useful for reducing or inhibiting viral infectivity, by contacting a cell (e.g. lymphocyte, nerve cell, connective tissue cell, muscle cell or hepatocyte) which is infected with a virus or is susceptible to infection with a virus, with an effective amount of the prototoxophore. The cells are cell lines an effective amount of the prototoxophore is also useful for ameliorating the severity of a viral infection in a subject, where the virus is selected from human immodeficiency virus (HIV), herpes simplex virus (FRV), rhinovirus and mativial contection or the prototoxophores of the invention are cells or bepatitis virus, by administering an effective amount of the present sequence represents an antivial contection or antivial prototoxophore is a secondary and present sequence represents and antiviral contection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             has
                                                                                                                               Hepatitis C; ds; viral prototoxophore; anti-viral; tumour; virus; infection; antitumour; toxophore; human immunodeficiency virus; HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Novel synthetic viral prototoxophore for treating viral infections, toxin molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  antiviral prototoxophore associated DNA sequence, this sequence is described as a recombinant NS3/NS4 fusion protein in example 1 of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  the invention although it is clearly not a protein sequence.
                                                                                       Anti-viral synthetic prototoxophore associated DNA sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Shepard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Example 1; Page 62; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                              26-APR-2002; 2002WO-US13223.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Neuteboom STC,
                                                                                                                                                                                                                                                                                                                                                                                                         27-APR-2001; 2001US-286893P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                    WO200287500-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cathers BE,
                                                                                                                                                                                                                                                                                                                 07-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          toxophore
```

```
WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                             P-PSDB; AAP9204
                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                        Houghton M,
                                                                                                                                                                                                                                                                                                     nepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     18
         (CHIR )
                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                           100 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 119
                                           79 GCCCAGCAGAAGAGGGCCTCCTAGGGTGCATAATCACCAGCCTAACTGGCCGGGACAAA 138
                                                                 160 AlaGlyHisAlavalGlyIlePheArgAlaAlavalCysThrArgGlyValAlaLysAla 179
                                                                                                                                                                                                                                              Combined open reading frames of the hepatitis C virus (HCV) cDNA in clones 14i 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c,
                GGTAGTGTGGTCATTGTGGGTAGGATCATTTTGTCCGGTAGTGGTAGTATCACGGCGTAC 78
                               AlaGinGinThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 39
                                                           AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPhcLeuAlaThr 59
                                                                                       CysileAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 79
                                                                                                                   ProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 99
                                                                                                                                                                                                                                                                        180 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 194
                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
Location/Qualifiers 3..5300 /*tag= a
                                                                                                                                                                                                                                                                                                            AAN92097 standard; DNA; 5300 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      870S-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87US-0122714
                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                        (updated)
                                                                                                                                                                                                                                                                                                                                                                    clones 14; 11b, 7f, 7
8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                       25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30-DEC-1987,
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-MAY-1988
26-0CT-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                   EP318216-A
                                                                                                                                                                                                                                                                                                                          AAN92097;
                                                           40
                                                                                       9
                                                                                                                   80
                                                                                                                                                              319
                                                                                                                                                                                                                                                   499
                                                                                                                                                                                                                                                                              559
                                                                                                                                                                                                                                                                                                    AAN92097
                a
                               ö
                                          a
                                                           ò
                                                                        g
                                                                                     õ
                                                                                                   g
                                                                                                                  ò
                                                                                                                              셤
                                                                                                                                            ö
                                                                                                                                                            g
                                                                                                                                                                             င်
                                                                                                                                                                                        셤
                                                                                                                                                                                                        å
                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                             a
```

```
CORP.) (tag a) extending through clones 14i, 11b, 7f, 7e, 8h 33c, 40b, 37b, 35 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c of hepatitis C vitus (HCV) ronA. In creating the composite sequence the following cheres of the considered. Clone 33c contains a sequence the following contains a sequence the following contains a sequence the following contains a well as in 5 other overlapping clones 40b and 37c. In clone 33c, as well as in 5 other overlapping clones, nuclectide #789 is a G. However, in clone 37b the corresponding nuclectide is an A. This heterogenity may have important ramifications for protein folding.

Contains the corresponding residue in clone 7e and in 3 other overlapping clones is an A. Therefore the residue in clone 7e and in 3 other overlapping clones is an A. Therefore the residue in clone 8f is represented as a T. C. than a G because the corresponding residue in clone 33f and in 2 other overlapping clones is a T. The 3' terminal sequence of clone 33f is coverlapping clones is a T. The 3' terminal sequence of clone 33f is coverlapping clones as A. The 3' terminal sequence of clone 33f and in 2 other overlapping clones is an A. The 3' terminal clone 33f and a contains and a clone 33f is and in 2 other overlapping clones is an A. The 3' terminal sequence of clone 31f and 2 other corresponding residue in clone 31f and 2 other overlapping clones is an A. The 3' terminal residue in clone 31f and 2 other corresponding artifacts have been omitted and instead clones is an A. The 3' terminal clones is a some contains and source of clones is an A. The 3' terminal sequence in clone 31f and 2 other corresponding artifacts have been omitted and instead clones is an A. The 3' terminal residue in clone 31f and 2 other corresponding sequences in non-5' terminal residue as a source of clone size and passive immuno-assay reagents and vaccines and to generate antibodies useful in hand title and passive immunotherapy for HCV infection/non-A, non-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                          is a double-stranded nucleotide sequence of the open reading frame
                                                                                                                                                                                                                                                                                                                                          Claim 3; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 5300 BP; 1047 A; 1606 C; 1515 G; 1130 T; 2 other;
                                                                                                                                                                                                             Purified hepatitis C virus
- and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-12 (1-195) x AAN92097 (1-5300)
Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2.15e-71
894.50
89.71%
85.78%
87.61%
Choo QL,
```

```
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                     286
                                                                                                                                                                                                                                 1047
                                                                                                                                                                                                                                                             1107
                                                                                                                                                                                                                                                                                        1167
                                                                                                                                                                                                                                                                                                                                 132
                                                                                                                                                                                                                                                                                                                                              1287
                                                                                                                                                                                                                                                                                                                                                            152
                                                                                                                                                                                                                                                                                                                                                                         1347
                                                                                                                                                                                                                                                                                                                                                                                       172
                                                                                                                                                                                                                                                                                                                                                                                                    1407
                                                                                                                                                                                                                                                                                                                                                                                                                  192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAN92103;
                                                                                                                                                                           927
                                                                                                                                                                                       32
                                                                                                                                                                                                                   52
                                                                                                                                                                                                                                                                           92
                                                                                                                                                             18
                                                                                                                                                                                                                                                                                                      112
                                                                                                                                                                                                                                                                                                                                                                                                                                1467
                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAN92103
                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                             BXHAXHXB
 8 \times 888888 \times 8
                                                                                                                                               셤
                                                                                                                                                                                                     a
                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                   ò
                                                                                                                                                             õ
                                                                                                                                                                          a
                                                                                                                                                                                       ò
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                      151
                                                                                                             152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The sequence shows the composite cDNA sequence derived from the aligned hepatitis C virus (HCV) cDNA's in clones 141, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c. The cDNA
       72 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                  ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                        Repatitis C virus; composite cDNA; probe; vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Fig. 26; 174pp; English.
                                                                                                                                                                                                                                                                                                    Location/Qualifiers 3..5360
                                                                                                                                                                                                                                                           Hepatitis C virus composite probe
                                                                                                                                                                                                            ВÞ
                                                                                                                                                                                                            CDNA; 5360
                                                                                                                                                                                                                                                                                                                                                                        870S-0122714.
870S-0139886.
880S-0161072.
880S-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                         Kuo G
                                                                                                                                                                                                                                                                                                                                                           88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                           (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                         (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                    ď
                                                                                                                                                                   MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                        Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI: 1989-215054/30
                                                                                                                                                                                                            AAN90327 standard;
                                                                                                                                                                                                                                                                                       Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                           18-NOV-1988;
                                                                                                                                                                                                                                       25-MAR-2003
11-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                         18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                               30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                GB2212511-A
                                                                                                                                                                                                                                                                                                                                             26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                             26-OCT-1988
                                                                                  132
                                                                                                                                                                  192
                                                                                                                                                                                                                         AAN90327;
                                                                                                                                                                                              RESULT 12
                                                                                                                                                                                                                                                                                                     Key
                                                                                                                                                                                                     AAN90327
                                                                                                                                                                                                           g
                                                                                             a
                                                                                                                                                                               g
             셤
                           ð
                                      g
                                                      ò
                                                                                 õ
                                                                                                             à
                                                                                                                           윱
                                                                                                                                       ŝ
                                                                                                                                                    a
                                                                                                                                                                   ð
```

```
CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArg 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                                                            CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71
                                                                                                                                                                                                                                                                                                                                                                                                                                            encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The cDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP90158, AAN90303-26, and AAN90328-36. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                  3 LysLysGlySerValValIleValGly-----ArgIleValLeuAsnGly----
                                                                                                                                    Sequence 5360 BP; 1060 A; 1622 C; 1532 G; 1145 T; 1 other;
                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                (1-5360)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ВP
                                                                                                                                                                                                                                                                                                                              US-09-965-594-12 (1-195) x AAN90327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAN92103 standard; DNA; 6905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATGAGGTCCCCG 1478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                              894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (updated)
                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
```

1619

151

131

```
Combined open reading frames of the hepatitis C virus (HCV) cDNAs from clones K9-1 through 15e.
                                                                                                                                                                      92 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                   112 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
  72 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                          132 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Purified hepatitis C virus
- and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 3; Figure 47-1 - 47-8; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers 3..7310
                                                                                                                                                                                                                                                                                                                                                                                            AAN92106 standard; DNA; 7310 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ပ်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87US-0122714.
87US-0139886.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Kuo
                                                                                                                                                                                                                                                                                                                                1740 ATGAGGTCCCCG 1751
                                                                                                                                                                                                                                                                                                                                                                                                                                               (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                  MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1989-159274/22.
P-PSDB; AAP92050.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EP318216-A.
                                                                                                                                                                                                                                                                                                                 192
                                                                                                                                                                                                                                                                                                                                                                                                                      AAN92106;
                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                   AAN92106
                          g
                                                                                                       õ
                                                                                                                                 a
                                                                                                                                                          ö
                                                                                                                                                                                 셤
                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1100 TIGCIGGCGCCCAICACGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAAIC 1259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1140 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                      It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 LysLysGlySerValValIleValGly----- 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -------AlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 other;
                          non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                           Purified hepatitis C virus and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                               Claim 3; Figure 32-1 - 32-7; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-12 (1-195) x AAN92103 (1-6905)
                                                                              Location/Qualifiers
                          Hepatitis C virus; HCV; non-A,
                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                     870S-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2.97e-71
894.50
89.71%
85.78%
87.61%
                                                                                                                                                                                  88EP-0310922
                                                                                                                                                                                                           87US-0122714
                                                                                                                                                                                                                                                                                                                              Kuo
                                                                                                       Ø
                                                                                           3..6905
/*tag=
12f through 15e
                                                                                                                                                                                                                                                                                                                            Choo QL,
                                                                                                                                                                                                                                                                                                                                                     WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                     CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                    virus
                                                                                                                                                                                                                                                                                                                                                                   P-PSDB; AAP92047
                                                                                                                                                                                                                                                                                                   (CHIR ) CHIRON
                                                                                                                                                                                                                                                            26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
                                                                                                                                                                                  18-NOV-1988;
                                                                                                                                                                                                                                                                                                                            Houghton M,
                                                    Hepatitis C
                                                                                                                                                          31-MAY-1989
                                                                                                                                                                                                           18-NOV-1987
30-DEC-1987
                                                                                                                                                                                                                        DEC-1987
                                                                                                                                                                                                                                      26-FEB-1988
                                                                                                                                                                                                                                                  06-MAY-1988
                                                                                                                                 EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
 clones
                                                                              Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
```

윤 õ g ò g ö

ò

(updated)

```
18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
    19-JUL-2001
             01-NOV-1989
                                                                                                                                            18-NOV-1988;
                                                                                                                                                                                                                                           Houghton M,
                                                                                                 GB2212511-A.
                                                                                                                       26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                   It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                            1725 TIGCIGGCGCCCATCACGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATC 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1965 GACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACT 2024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1CGGGGGGTCCGCTGTTGTGCCCCGCGGGCACGCCGTGGGCATATTTAGGGCCGCGGTG 2204
                                                                                                                                                                                                                                                                                                                                         151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       131
                                                                                                                                                                                                                                                                                                                                                                                                                                                           AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 191
                                                                                                                                                                                                                                                                                                                              51
                                                                                                                                                                                                                                                                                                                                                                                                                    91
                                                                                                                                                                                                                                                                                                                                                                        52 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 71
                                                                                                                                                                                                                                                                                    ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                   72 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 132 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                           ----ArgileValLeuAsnGly--
                                                                                                            Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 4 other;
                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                           LysLysGlySerValValIleValGly----
                                                                                                                                                                                                                      US-09-965-594-12 (1-195) x AAN92106 (1-7310)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAN90336 standard; DNA; 7310
                                                                                                                                          3.19e-71
894.50
89.71%
85.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2265 ATGAGGTCCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (updated)
                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2145
                                                                                                                                                                                                                                           ٣
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      172
                                                                                                                                                                                                                                                                                    18
                                                                                                                                                                                                                                                                                                                              32
                                                                                                                                                                                                                                                                                                                                                                                                                                                             92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAN90336;
                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAN90336
ID AAN9
  g
                                                                                                                                                                                                                                                                                                         윱
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
1785 ACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATTGTGTCAACT 1844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    72 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes artigons which react with antibodies in patients with non-hon-B hepatitis (NAMBH). The cDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP90208, and AAN90303-35.

(N.B. This record was resubmitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       52 AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 LysLysGlySerValValIleValGly----ArgIleValLeuAsnGly---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-12 (1-195) x AAN90336 (1-7310)
                                                   Composite hepatitis C virus (HCV) cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; fig 47; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                              87US-0122714
87US-0139886
                                                                                                                                                                                                                                                                                                                                                                                                                                     88US-0161072
88US-0263584
                                                                                                                                                                                                                                                                                                                              88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (CHIR ) CHIRON CORPORATION
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               894.50
89.718
85.788
87.618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1989-215054/30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P-PSDB; AAP90288
                                                                                                                                                             Pan troglodytes.
```

```
1905 GCCGGAACGAGGACCATCGCGTCACCCAAGGTCTIGTCATCCAGATGTATACCAATGTA 1964
                                                                                  11111111111
2265 ATGAGGTCCCCG 2276
                                                                              192 MetArgSerPro 195
                                         4 5 8
                                                              03
05
                                                                            ò
                           g o
q
       Š
             g
                     ò
```

Search completed: August 30, 2003, 19:47:47 Job time : 190.082 secs

BF304699 601888252

Description

29: gb_gss2:*

```
BF304699 984 bp mRNA linear EST 21-NOV-2000 601888252F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4122276 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 984)
         Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALIGNMENTS
                                                    SUMMARIES
                                                                                                                                    BQ926101
BI198486
BI0881847
BJ001625
BJ024121
BJ016176
BF203316
BF631437
BM915803
                                                                                                                                                                                                                                                   BF182274
BF182273
BU169585
AW785806
BC420860
CC221189
BM803824
AU127824
BQ894657
BE289911
                                                                                                                                                                                                                                                                                                                                                        CB993468
B0709745
CD262790
BG387051
BG582842
AL555424
BQ928211
BI196460
CR21699
AU1251699
BU540812
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BX451426
CD105862
BQ672290
BU556872
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BM548430
BE622016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BQ919246
CB286751
                                                                                                      BF304699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BQ673186
                                                                                                                           BU148820
                                                                                                                                                                                                                                        BF863244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BF304699.1 GI:11251586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (human)
Homo sapiens
                                                                                   DB
                                                                                   Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               mRNA sequence.
BF304699
                                                                        Query
                                                                                                                                                                                                                                                                                                                                                                                     93.5
93.5
93.5
93.5
                                                                                 Score
                                                                                                                                                                                                                                                             95.5
95.5
95
                                                                                                                                                                                                                                                                                                                                                                                                                                         103.5
103
101.5
100.5
                                                                                                                                                                                                98.5
98
                                                                                                      106
                                                                                                                                                                                                                                                                                                                                                                            93.
                                                                                                                                                                                                                   97.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BF304699/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
                                                                         Result
                                                                                   .
8
                                                                                                                                                                                                                                                                                                                     (without alignments)
2506.388 Million cell updates/sec
                                                                                                              1021
1 MKKKGSVVIVGRIVLNGAYA......VAKAVDFIPVESLETTMRSP 195
                                                           August 30, 2003, 19:20:43 ; Search time 1890.92 Seconds
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd

    nucleic search, using frame_plus_p2n model

                                                                                                                                                                                                          22781392 seqs, 12152238056 residues
                                                                                                                                                                                                                             hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                       0.0
0.0
0.0
0.0
                                                                                                                                           BLOSUM62
Xgapop 10.0, Xgapext
Ygapop 10.0, Ygapext
Fgapop 6.0, Fgapext
Delop 6.0, Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_gss_fun:*
em_gss_mam:*
em_gss_mus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_gss_rod:*
em_gss_phg:*
em_gss_vrl:*
gb_gssl:*
                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_gss_inv:*
em_gss_pln:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gb_est2:*
gb_htc:*
gb_est3:*
gb_est4:*
gb_est5:*
em_estfon:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_gss_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_gss_vrt:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_gss_pro:*
                                                                                                    US-09-965-594-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_esthum:*
em_estin:*
em_estmu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       em_estro:*
em_htc:*
gb_est1:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_estov:*
em_estpl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_estba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                  EST: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Total number of
                                                                                                    Title:
Perfect score:
                                                                                                                                            Scoring table:
                                         OM protein
                                                                                                                          Sequence:
                                                                                                                                                                                                           Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Catabase
                                                            Run on:
```

B0926101 AGENCOURT B198486 602760491 B0981847 AGENCOURT BJ09186 602760491 B0981847 AGENCOURT BJ0016176 BJ016176 BF203316 601865914 AK080545 Mus muscu BF31437 HVS/MED001 BH915803 AGENCOURT BF863244 96304202 BF3307233 601891502 BHJ85806 6117260 MA BM9186806 6117260 MA BM9186806 117260 MA BM918274 AGENCOURT BM9894657 AGENCOURT BM9894657 AGENCOURT BM994657 AGENCOURT BM9094657 AGENCOURT BM9094657 AGENCOURT BM9094657 AGENCOURT BM9094657 AGENCOURT BM9094657 AGENCOURT BM909466 G02454179 BC342842 AGENCOURT BM90941426 BM4019XE BM90941426 BM4019XE BM90941426 BM451426 BM556872 AGENCOURT BM556872 AGENCOURT BM556872 AGENCOURT BM556872 AGENCOURT BM556872 AGENCOURT BM556873 AGENCOURT BM556873 AGENCOURT BM556873 AGENCOURT BM558143 AGENCOURT BM56873186 AGENCOURT BM56873186 AGENCOURT BM56873186 AGENCOURT

BE622016 601440668 BQ919246 AGENCOURT CB286751 CMD45_C08

```
106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
9
                                                                           REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                1. .984

/organism="Homo sapiens"
/organism="Homo sapiens"
/mol_type="mrNA"
/db_xref="taxon:9606"
/clone="!MAGE:412276"
/clone="!MAGE:412276"
/clone="!MAGE:412276"
/clone="!MAGE:417"
/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BQ892487 1199 bp mRNA linear EST 16-AUG-2002 AGENCOURT_8417538 Lupski_sympathetic_trunk Homo sapiens cDNA clone IMAGE:6192708 5', mRNA sequence.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Genc Collection (MGC)
Lupublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs remail.nih.gov
Tissue Procurement: ATC
Tissue Procurement: ATC
Tissue Procurement: ATC
ToNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1005 row: g column: 13
High quality sequence stop: 646.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    117 LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98 TrpProAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       137 GlySerLeuLeuSerProArgPro--------IleSerTyrLeuLysGlySer 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     553 GGCGCCTCTTGTGGGGAAGACCTCGATGGTGTCCAAGCTCGCGCTGCTGTACTGGAAGT 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   152 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ------TICCAGCCCCGGGG 455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 598 ------ACCAGGCACGCAACATACATGCAAGGAGACGTGGT---TCCCGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             984
33
24
22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        493 CGCACGCTCCGGTCAGGTGCAGC-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-12 (1-195) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BQ892487
BQ892487.1 GI:22284501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        454 TGCGCCCGAGGA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      172 CysThrArgGly 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
   AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BQ892487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
/dolone_lib="Lupski.sympathetic_trunk"
/flone_lib="Lupski.sympathetic_trunk"
/flone="Vector: pcMv-SPORT6 (Life Technologies); Site_l:
NOTI: Site_2: Sali; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCG-3' and
5'-TCGACCCACGCGTCGG-3' and
5'-TCGACCCACGCGTCGG-3' and
1' kD for average insert length 1.9 kb. This is a primary library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor college of medicine); available through Life
                                                                         NIH'MGC http://mgc.nci.nih.gov/.

National institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: GapbS-rémail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Preparation: Life I.M.A.G.E. Consortium/LLNL at:
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC column: 13
High quality sequence start: 57

Corganism="Homo sapiens"
//Organism="Homo sapiens"
//Corganism="Homo sapiens"
//Colone="Homo sapiens"
//Colone="Homo sapiens"
//Colone="Homo sapiens"
//Colone="Homo sapiens"
//Colone="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        484 TGGGATCCATTTTTAATAAAGGGTGCTCTGTTAATCATGGCGCCCACGGCCCCGCTGATA 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   86 GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArg 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     574 ---TGIGCIGCIGCACACCCCCATGACCGATGTTGGGCTTATGTGGAAACGGCGAG 630
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ----ArgArgGlyAspSerArgGlySerLeuLeu-- 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    141 ....-SerProArgProlleSerTyrLeuLys-------GlySerSerGl 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            66 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValile 85
Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Buteleostomi,
Mammalia, Butheria, Primates, Catarrhini, Hominidae, Homo.

    (bases 1 to 1199)

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      117 LeuTyr-LeuValThr------ArgHisAlaAspValIleProValArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1199
41
17
53
44
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              211 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-12 (1-195) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Technologies."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /sex-"male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      103.50
37.428
26.458
10.148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 362 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
```

a õ a å g ò 셤 ö

ò

q õ ద g ö δ

à

```
/clone="InAGE:4895638"
/tissue_type="neuroblastoma"
/lab_host="DH108 (phage-resistant)"
/clone_lib="NH1M_GC_19"
/clone_lib="NH1M_GC_19"
/clone_lib="NH1M_GC_19"
/clone_lib="NH1M_GC_19"
/note="Organ: brain; Vector: pOTB7; Site_l: XhoI; Site_2:
ECORI; CDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sitea using the following 5'
adaptor: GGCACAGG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NHLMGC Library."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  554 CCCACTITGICICCCAGGICITGAICAGCTGCACATCACTCAGGAITTCAAIGGIGCCCCI 495
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          314 CAAAGCCCACATICCCCCCGGATATGAGGCAGCAGT-TTCTCCAGAGCTGGGTTGTTTCC 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   255 AGGIGCCCTCGGAIGGCCTIGCGCATCAIGGTGTICTIGCCCAICAGCCACACACCTC 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     791 CGGTTGATGATAGAATGGGGTTACTGATTGCAACAGTTGAGTACCCAATCCTGCAGACAG 732
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89 ThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThr 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIlePro 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           494 GGAGATTTTAGTGGTGATACCTAAAGCCTGGAAAAAGGAGGTCTTCTCGGCCCGAGACCA 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----SerTyrLeuLysGlySerSerGlyGly----- 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    434 GIGITCIGGGCIGGCACAGIGACIICAGCAIGGGGCAAIGGCACCAGCACGGGICAGCAG 375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         374 CTGGCACCTTATTGGCCAGCAACATGTCCCTGATCTCAGTGAGGTCCTCCTTGGTGAACA 315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -------AlaGlyHisAlaValGlyIlePhe---ArgAlaAlaVal----- 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         52 AlaAlaGlnThrPhe---LeuAlaThrCys-----IleAsnGlyValCysTrpThrVal 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12 ArgileValLeuAsnGlyAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --LeuGluThrThrMetArgSerPro 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   671 ATATCAAGCACTTTCAGGGTTTGTAGAIGCTGCCAITTGTCGAACACCTGCTGG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         575 AGCAGCGTG-------GCTTCGCTGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              129 ValArgArgArgGlyAspSer-----ArgGlySerLeuLeuSerProArgPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----ProLeuLeuCysPro------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ---CysThrArgGlyValAlaLysAlaValAspPhelleProValGluSer-----
                                                                                                                                                                                                                                                                                                                                                                                                                       844
50
35
77
11
                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-12 (1-195) x BI198486 (1-844)
                          /db_xref="taxon:9606"
       /mol_type-"mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                  12.1
100.50
35.56%
20.92%
9.84%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      145 Ile-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                             203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 188
                                                                                                                                                                                                                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                       Pred, No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BI198486 844 bp mRNA linear EST 10-JUL-2001 602760491F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4895638 5',
                                                                                                                                                                                                                                  1381 AGAGGGTGTGTTCANCGGTCAGGACAGGTCGCCGCACACTCGACGCGTCGGCCAGAGACT 1322
                                                                                                                                                                                                                                                                                                                        1321 TGTCGGGGGCGCGTTGGCGCATACCCCGGGTCGGATCGAGGTCAGCGCCGCTTGTATACA 1262
                                                                                                                                                                                                                                                                                                                                                                                                                  1261 GAGGGGAAA-----CAG 1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1249 GGGGTA---TGGTTATCACGGGCTGGGGCAGGTACT-----TCCCCTAAAGCG 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1144 AA---GAGAGGCACCIGACGCCCTCCCCTGTGGGCNTGTCGATAATAACAAATGTGCAG 1088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1087 GGCACGGIGAIGIGGITACIACGGGGCAGACCGGCICCAACGGCGCTCTCIAACAGACGC 1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   82 lyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProG 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              102 InGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrA 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ....-SerArgGlySer-----LeuLeuSerProArgProIleSerTyrLeuLysG 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MG.E. Consortium/LLNL at: image.llnl.gov Plate: LLCM1781 row: f column: 23 High quality sequence stop: 794.

Location/Qualifiers
1. .844
                                                                                                                                                                                     LysGlySerValValIleValGlyArglleValLeuAsnGlyAlaTyrAlaGlnGlnThr
                                                                                                                                                                                                                                                                               ---ArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLysAsnGlnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                             63 GlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys--G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 (bases 1 to 844)
NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                     43 GluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 rgHisAlaAspValIleProValArgArgArgGlyAsp-----
    46
15
15
68
37
                       Conservative:
Mismatches:
Indels:
    Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism-"Homo sapiens"
                                                                                                                                          (1-1403)
                                                                                                                                       US-09-965-594-12 (1-195) x BQ926101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mRNA sequence.
BI198486
BI198486.1 GI:14653507
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    150 lySerSerGlyGly 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 967 GAGAAGGCGGCGGT 954
101.50
36.758
27.718
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unpublished
                Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              135
                                                                                                                                                                                                                                                                               24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
BI198486/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
```

S

```
BJ001625
BJ001625 MF01SSA CDNA Oryzias latipes CDNA clone MF01SSA025C02 5',
                                                                                                      501
                                                                                                                                                         -----TrpProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySer 114
                                                                                                                                                                                                             468
                                                                                                                                                                                                                                                             134
                                                                                                                                                                                                                                                                                                              423
                                                                                                                                                                                                                                                                                                                                                                 154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Acthopoterygli; Neoperygli; Teleostel; Biteleostel; Neoteleostel; Acanthomorpha; Acanthopoterygli; Percomorpha; Atherinomorpha; Beloniformes; Adrianichthyldae; Oryzlinae; Oryzlas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 (bases 1 to 615)
Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                                                                                                                    -----AAGGGAGGGGGGGGGGA 384
39 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 58
                                                     97
                                                   78 AlaSerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGly
                                                                                                                                                                                                                                                          115 SerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAsp
                                                                                                                                                                                                                                                                                              135 SerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /sex-mixture of female and male."
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/lone_lib="MFDISSA CDNA"
166 c 165 g 144 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          615
42
9
50
50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
                                                                                                                                                                                       500 GTATAACGGTGGCCGGGGCCTCGTGGTTCTCGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
                                                                                                      GGGGGGGGAGGAGGTGGGTTAATT-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism-"Oryzias latipes"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   [ndels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Oryzias latipes (Japanese medaka)
Oryzias latipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /mol_type="mRNA"
/strain="Hd-rR"
/db_xref="taxon:8090"
/clone="MF01SSA025C02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-12 (1-195) x BJ001625 (1-615)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .615
                                                                                                                                                                                                                                                                                                                                                                                                                  422 ACGGAGGCGGGATTGTATCG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BJ001625.1 GI:17364516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             111.3
99.00
33.77%
27.81%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             =
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mRNA sequence.
BJ001625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     155 Pro 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        383 CCG 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EST
                                                                                                      533
                                                                                                                                                         86
                                                                                                                                                                                                                                                                                                              467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Oucry Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
BJ001625/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ..
No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                   ò
                                                                                                   g
                                                                                                                                                         ò
                                                                                                                                                                                                  D
                                                                                                                                                                                                                                                          á
                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1. .1204
/organism="Homo sapiens"
/mol_type="mrnn"
/mol_type="mrnn"
/db_xref="taxon:9606"
/db_xref="taxon:9606"
/dbost="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mnnorganism="mn
                                                                              HUNSHIN4/ 1204 DP MRNA linear EST 16-AUG-2002
AGENCOURT_8712410 NIH_MGC_112 Homo sapiens CDNA clone IMAGE:6295246
5', MRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GTGCCAAGGCGTGTGCTCACGGGGCAGTATATGGCGCCAAACCGTGGGGTGCGGAACCGCC 714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 IleAsnGlyValCysTrp----ThrValTyrHisGlyAlaGlyThrArgThrIle 77
                                                                                                                                                                                                                                                                                                        Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo. I (basea 1 to 1204)
NIH-MGC http://mgc.noi.nih.gov/.
195 CCGCGAAGGGACATGCGGATCTGCTGCTTGCTTGGAGCCCACATTGTCTGCTCCC 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDNA Library Preparation: Rubin Laboratory
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MCC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://mage.lnl.gov
Plate: LLCM2501 row: c column: 23
High quality sequence stop: 322.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26 -------GluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLysAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 GlnValGluGlyGluValGLnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10 ValGlyArgIleValLeuAsnGlyAlaTyrAlaGlnGlnThrArgGly--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1204
41
17
17
58
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-12 (1-195) x BQB81847 (1-1204)
                                                                                                                                                                                                       80881847.1 GI:22273855
                                                                                                                                                                                                                                                          Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19
100.50
36.028
25.478
9.848
                                                                                                                                                                                                                                                                                                                                                                                                                                          Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                            BQ881847
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                   RESULT 6
BQ881847/c
                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
∑
                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SASE COUNT
                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORIGIN
```

g

ò 셤

õ g ò

Best Local Similarity: 27.81% Mismatches:	12 95) x BJ024121 (1-	Oy 39 LysAsnGlnValGluGlyGluValGlnIleValSerThrA		79 SerProLys 344 GCTCTGCGGCGAG	Qy 94 AspLeuValGlyTrpProAlaProGlnGlySerArgSerI	Oy 110CysThrCysGlySerSerAspLeuTyrLeuV	Qy 123	Oy 136 ArgGlySerLeuleuSerProArgProlles	/db_xref="taxon" /clone="MF01SSA025C02" /sex="mixture of female and male" /tissue_type="whole embryo" /dev_stage="segmentation stage 20
511 AAAAATGACGTAGAACCAAAAAGACACAGATCCAACCACATGTTCTGGTTCTACGGGCT 452	59 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 78 	79 SerProLysGlyProVallleGlnMetTyrThrAsnValAspLys 93 ::: ::: 409 GCTCTGGGGGAGGAGGAGGAGGGCTAGGTAGGGGGGGGGG	AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro	CTGCAGGGATGCGGATGTCTGCT	123	136 ArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSer 152 223 AGAGGCGACAGGAGGTGGAGGTCTCTGCAGGACGGGGACA 164	153 GlyGlyProLeuLeuCysProalaGlyHisala 163 	SULT 8 CUSA 121 BJ024121 CESSION BJ024121 BJ024121 BJ024121 BJ024121 CESSION BJ024121 CESSION BJ024121 CESSION BJ024121 CESSION BJ024121 CESSION BJ024121 ACTIONE MPREASOR CHORDARS LATIONS CONTROLED	Alignment Scores: 12 Length: 643 Pred. No.: 99.00 Matches: 42 Score: 33.77% Conservative: 9
යි	oy O	oy fe	ි ඊ යි	Qy	Oy Pp	OY Db	Qy Db	RESULT 8 1024121 10CUS DEFINITION ACCESSION VERYWORDS SOURCE ORGANISM TITLE JOURNAL COMMENT COMMENT FEATURES SOUICE BASE COUNT ORIGIN	Alignment Pred. No.: Score: Percent Si

```
linear EST 05-DEC-2001
NA clone MF01SSA025C02 3',
                                                                                                                                                                                                                                                                                    eSerTyrLeuLysGlySerSer 152
||| || || || || || AGGACGTCTTCACGGCGGATCA 589
                                                                                                                                                                                                                                 rLeuThrPro----- 109
                                                                                                                                                                                                                                                                                                                                       ovalArgArgArgGlyAspSer 135
||||||
CGCTGGTCTGACGGGTGATGGG 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Vertebrata; Euteleostomi;
Euteleostei; Neoteleostei;
pha; Atherinomorpha;
le; Oryzlas.
                                                                                                                                                                                                                                                     AGCGTCACCCGGAGCTGTAGG 418
                                                                                                   ACATGITCIGGITCIACGGGCT 301
                                                                                                                                               T., Jindo, T. and Takeda, H.
                                                                                                                             yAlaGlyThrArgThrIleAla 78
                                                                                                                                                                                 nMetTyrThrAsnValAspLys 93
                                                                          ralaalaGlnThrPheLeuala 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         20 - 25
50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .
```

```
// Abdustance of the control of contr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Indh, M., Konno, H., Okazaki, Y., Muramafsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected CDMAs to prepare full-length cDNA libraries for rapid discovery of new genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer-AspLeuTy 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            158 ysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaL 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    477 GTCCTGCAGTGCTGCTGTAGGTGCTGTCCCGTCAGCTTTTTGTGATGGTGGTGACGCTG 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARUBUD45
Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:A730082L10 product:weakly similar to zinc
finger protein (fragment) [Mus musculus], full insert sequence.
AK080545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               118 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              591 GACTCTGAGTGCAGGAGCGGACAGAGGGGATCGGAAGCGCGGGGATGCTGATGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             642 CCGCCTCCTCCTCCTCGGGCA-----TGAGAATGTGGAGCTCAGAGATGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 138 rLeuLeuSer-ProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                32
111
32
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
                             /organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-12 (1-195) x BF203316 (1-961)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Carninci, P. and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AK080545.1 GI:26348600
HTC; CAP trapper.
Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     52.44%
39.02%
9.65%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             178 ys 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AA 416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0349636
                                                                                                                                                                                                                                                                                                                                                                                                                                                              230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
No
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 11
AK080545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BF203316 961 bp mRNA linear EST 06-NOV-2000 601865914F1 NIH_MGC_17 Homo sapiens CDNA clone IMAGE:4098578 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     469
                                                                                                                                                                                                                                                                                                                                                                                                                                                            301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -----HisAlaAspValIleProValArgArgArgGlyAspSer 135
                                                                                                                                                                                                                                                                                                                                                                                                      28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 961)
NIH-MGC http://mgc.nci.nih.gov/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Contact: Robert Strausberg, Ph.D.

Email: Cgapbs-rémail.nih.gov

Tissue Procurement: ATCS

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incytte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCMSG5 row: 1 column: 03

High quality sequence stop: 637.

Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                   39 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 59 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          419 CTGCAGGGATGCGGATGTGCCTCTGCT-----TTGGTTCCTCCTCCTGGATCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  79 SerProLys------GlyProValileGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   380 ------ACTCGTGGAGGACGAAGAGCGTCACCCCGGAGCTGTAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        136 ArgGlySerLeuLeuSerProArg-----ProlleSerTyrLeuLysGlySerSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              530 AGAGGCCGGACAGCAGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                         242 AAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        344 GCTCTGCGGCGGCGAGGAGCTCCTGGGCCAGTTGTG--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            110 ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                470 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGG
                                                                                                                                         754
90
50
50
                                                                                                                                                                                              Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       153 GlyGlyProLeuLeuCysProAlaGlyHisAla 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            590 GGAGGACCGACTCGCTGCAGAGCCTCTGCTGCA 622
                             u
                                                                                                                                            Length:
Matches:
                          198
/clone_lib="MF015SA cDNA"
                                                                                                                                                                                                                                                                                                                                             US-09-965-594-12 (1-195) x BJ016176 (1-754)
                                5
                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BF203316.1 GI:11096902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens (human)
                                                                                                                                         14.6
99.00
33.778
27.818
                             181 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mRNA sequence.
BF203316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo saptens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unpublished
                                                                                                                                                                                                                                  Similarity:
                             ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          123 ......
                          194
                                                                                                                                                                                                 Percent Similarity:
                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                          Query Match:
ORIGIN
                                                                                                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BF203316/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
```

g

ò g à

a

ò 셤 ò a

ò

ò

a ð 용 a

ö

138

158

HTC 05-DEC-2002

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

TITLE

```
/protein_id="bac37940.1"
/db_xref="G1:26348601"
/db_xref="G1:26348601"
/translation="bsclpaaspcrsiltprcdcffikeklsaaravgpgpsvafgvs
Trvgadadafgcgrrygacrrseglclsrkprrrghvppvgphvyglygptyggpspage
AQAAGRAPQOVPHPGRRHGTVVPPQCAAGLLPALAARQVPVGPVRGREGPRRAPRHS
PRPVPTALGFSFGGGGPAPPLLAPANGRSVGLAL"
                                                                                                                                                                                                                                                                                                                                                                                                                                          /tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    372 GTGGTG-----CCTCCTCAAGGA 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           124 AlaAspValIleProValArg---ArgArgGlyAspSerArgGlySerLeuLeuSerPro 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             143 ArgProlle------SerTyrLeuLysGlySerSerGlyGlyProLeu 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 486 AAGCCGGTTCCTACAGCCTTGGGTTTCTCGTTTGGGCAGGGTGGGCCTGCTCCTCCCCTC 545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    157 LeuCysProAla---GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGly 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         104 SerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHis 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ---CGGCCACATGGAACA 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           176 ValAlaLysAlaVal-----AspPheIleProValGluSerLeuGluThrThrMet 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          603 GACTICAGGGIGGIIGGCICIGAGCIGACACCAAGCICCICICGGCAGGIACACII 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                44 GlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    303 GGAGAAGCGCAGCTGGACGCGCGCGCAG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 ValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84 ValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note="unnamed protein product; putative weakly similar to zinc finger protein (fragment) musculus) (PIRI148722, evidence: FASTY, 50.7%ID, 757.68length, match=601)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1141
46
11
52
50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                        /db_xref="FANTOM_DB:A730082L10"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ىد
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 336 ----CAAGTGCCGCATCCGCCTGGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                228
                 URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels:
                                                                                                                                                                                        'organism-"Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-12 (1-195) x AK080545 (1-1141)
                                                                                                                                                                                                                                                                                                                                                           /db_xref="taxon:10090"
/clone="A730082L10"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                353 g
                                                                                                                                                                                                                                /mol_type="mRNA"
/strain="C57BL/6J"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'note-"putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note-"putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31
98.00
35.85%
28.93%
9.60%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                316 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                244 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  polyA_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyA_site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDS
                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                              Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Mishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Ravai,J., Oxazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN Integrated sequence analysis (RISA) system--384-format Genome Res. 10 (11), 1757-1771 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Arakawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Salto, T., Ostazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Salto, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Rubol, P., Lewis, S., Matsuo, Y., Mikaido, I., Fesole, G., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Mikaido, I., Fesole, G., Tomita, M., Wanner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wanshizaki, Y., Banda, Y., Kawaji, H., Kohtsuki, S., Manda, M., Hasseqawa, Y., Kawaji, H., Kohtsuki, S., Manda, M., Manda
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

L Nature 420, 563-573 (2002)

E ( (bases I to 1141)

Radachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanaqaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramotto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kaqawa, I., Kasukawa, T., Kotoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Nakanura, M., Nomura, K., Numazaki, R., Ohno, H., Okazaki, Y., Saito, R., Sattoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Saszume, N., Sano, H., Saszume, N., Sano, H., Taqami, M., Taqawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tanak
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted Summission of Parker 2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Physical and Chemical Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi'ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-reségsc.riken,go.jp,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Division of Experimental Animal Research in Riken contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The FANTOM Consortium and the RIKEN Genome Exploration Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               URL: http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222.
Fax:81-45-503-9216)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              prepare mouse tissues.
Please visit our web site for further details.
Genome Rcs. 10 (10), 1617-1630 (2000)
20499374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
REFERENCE
AUTHORS
```

TITLE JOORNAL MEDLINE PUBMED REFERENCE AUTHORS

TITLE

485

TITLE

COMMENT

g

LOCUS

```
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                               29
                                                                                                                                                                                                                                                                                                                                                                                                                                                98
                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
BM915803/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Locus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /isou-rype="Seedling shoot"
/lab_host="TJC121"
/lab_host="TJC121"
/clone_lib="Hordeum vulgare seedling shoot EST library
/clone_lib="Hordeum vulgare seedling shoot EST library
/rote="Vector: lambdaZAP; Site_l: EcoR]; Site_2: Xhol;
Seeds were surface sterilized then germinated under axenic
conditions in the dark at room temperature on filter paper
with water, nystatin and cefotaxime in covered
crystallization dishes. Five-day old seedlings were
incubated at 90 RH for 24 hr. Shoots were then harvested,
total RNA was prepared, poly(A) RNA was multied, one
primary unamplified cDNA library was made, 600000 pfu were
primary unamplified cDNA library was made, 600000 pfu were
in vivo excised to give pBluescript SK(-) cDNA phagemids.
These steps were performed in the TJ Close laboratory at
the University of California, Riverside (Choi, Close,
Penton). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
above. For more details on library preparation and
                                                                                                                                                                                                                                                                                                                   Trittceae; Hordeum.

1 (bases 1 to 779)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton Development of a genetically and physically anchored EST resource for barley genomics: Morex drought-stressed seedling shoot cDNA
                           779 bp mRNA linear EST 22-OCT-2001 HVSMEb0015P05f Hordeum vulgare seedling shoot EST library archae stress) Hordeum vulgare subsp. vulgare cDNA BF631437
                                                                                                                                                                                                                          Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryora: Vuidiplanta: Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta: Liliopsida; Poales; Poaceae; Pooldeae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)* 216 216 311 g 105 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unpublished
On Dec 19, 2000 this sequence version replaced gi:11895595.
Contact: Wing RA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism="Hordeum vulgare subsp. vulgare"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /db_xref-"taxon:112509"
/clone="HVSMEb0015P05f"
/tissue_type="Seedling shoot"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Clemson University Genomics Institute
Clemson University
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Seg primer: AATTAACCCTCACTAAAGGG
High quality sequence stop: 588.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sequence analysis see
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /mol_type="mRNA"
/cultivar="Morex"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Email: rwing@clemson.edu
Total hq bases = 455
                                                                                                                                                                            BF631437.2 GI:13092107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   library
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
ORIGIN
                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
COMMENT
RESULT 12
BF631437
                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
```

```
AGENCORT_6639455 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482056 BM915803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    289 CCGCGGCCGTTGCCGGGGCACTATACG-TCGCCGGCGCGCGCACGACGAGGAGAAGAAGAAGGCG 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                109 CCCTGTTCTGCCAGCTGGCCGCGGTGGACGCGCGCGCAGGGGAAGCGGCTCGTGGTCG 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              116 -AspLeuTyrLeuValThrArg-----HisAlaAspValIleProVal-----ArgAr 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         131 gArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySe 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             348 CICCGGICGGCGCICGCGIACGACCCGGACTCCGACGCGIGGGCCGCGCTCCCGGACAIG 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa; Chordata; Cranlata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. I (bases 1 to 1146)
NIH-WGC http://mgc.nci.nih.gov/.
                                                                                                                                                                                                                                      ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100 AlaProGlnGlySerArgSer------LeuThrProCysThrCysGlySerSer
                                                                                                                                                                                                         GlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrp
                                                                                                                                                                                                                                                                                                    ThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIle---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDNA Library Preparation: Rubin Laboratory
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MG colone distribution information of
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov i column: 01
Plate: LLCM2007 row: i column: 01
High quality sequence start: 6
High quality sequence start: 6
Location/Qualifiers
779
47
14
63
30
6
                                           Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         408 GCGCAGGACGCGACGACGCGCGCGGGCTCTGCGT 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           163 laValGlyIlePheArgAlaAlaValCySThrArg 174
  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Contact: Robert Strausberg, Ph.D.
                                                                                                                     Gaps:
                                                                                                                                                              US-09-965-594-12 (1-195) x BF631437 (1-779)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Email: cgapbs-rémail.nih.gov
Tissue Procurement: DCTD/DTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BM915803.1 GI:19366182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens (human)
21.3
97.50
39.878
30.728
9.558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unpublished
                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
```

```
::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-12 (1-195) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Email: chauser@duke.edu.
Location/Qualifiers
                                                                             Chlamydomonas reinhardtii
                                                                                                                                                                                                                                                                                                                                                 Contact: Charles Hauser
DCMB Box 91000
Duke University
Durham, NC 27708-1000
Tel: 919 613 8159
Fax: 919 613 8177
                            BF863244.1 GI:12253388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26
96.00
40.71%
28.32%
9.40%
                                                                                                                                                                                                                                                                                                                               Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       173 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COUNT
                                                                                                  ORGANISM
                         VERSION
KEYWORDS
SOURCE
     ACCESSION
                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                        AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                      TITLE
                                                                                                                                                                                                                                                                                                                                                         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
               // Ab_xref="taxon:9606"
// Ab_xref="taxon:9606"
// Ab_xref="taxon:9606"
// Ab_xref="taxon:9606"
// Ab_nost="taxon:9606"
// Ab_nost="taxon:9606"
// Ab_nost="bul00 (phage-resistant)"
// Ab_nost="bul00 (phage-resistant)
// Ab_nost="bul00 (phag
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BF863244 C. reinhardtii CC-1690, Stress condition I, normalized , Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gHisAlaAspValileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerPr 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                833 ......CGGGAGGGCGCGGGCGGGCCCGGTAGTGTGGCCCCGG 796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142 oArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHi 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPh 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -------GGIGGGAIGTIGTGG-----TAICACTICCCGGCGCGGGGGG 958
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103 -GlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrAr 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      eLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHis------GlyAl 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   17 GlyAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSer-GlnThrGl 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36 yargaspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    aGlyThr------LysGlyProvalll
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  957 AGGTACGTGAGCGGAGGGCGCCGCGTCGGGGCGCCGGGGCGCGCGGGGGCCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 eGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGln-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             867 GGGGGCCAGACTTGCTTGTCGTGTTTTCTGTGGG-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels:
/organism-"Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-12 (1-195) x BM915803 (1-1146)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    182 e---IleProval 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      690 TCGCGTACCTTTG 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38.9
97.00
34.05%
26.49%
9.50%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ..
02
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BF863244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
```

```
/db_xref="taxon:3055"
/db_xref="taxon:3055"
/db_xref="taxon:3055"
/clone_llb="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambdd 2ap II"
/note="Vector: paluescript II SK-; Site_I: ECORI; Site_2:
XhOI; This library, constructed by John Davies and Jeffrey
McDermott, combines conNas from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, lhr, 4hr), TAP-S (30 min,
lhr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, lhr,
'4hr) and NH4 to NO3 (30min, lhr, 4hr). POLYA mRNA was
purified from each sample, pooled and cDNA synthesized.
The cDNA was directionally cloned into lambda Zap II
(Stratagene) in the ECORI (5') and XhORI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with Exassist (Stratagene)
phage. The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     125 AspVallle----- 136----- 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --- GlySerLeuLeuSerProArgProIleSerTyrLeu 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         285 AGAAGCCCGACCGCTGGTCCTCTGGGGTCATCCGCATGCTATGCAATCTCCCGCTATCAG 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        85 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                105 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             345 GAGATCATITGCATGTGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTTATAA 404
chiamydomonas reinhardiii
chiamydomonas reinhardiii
chiamydomonas reinhardiii
chiamydomonadaceae: Chlamydomonas.
1 (bases 1 to 701)
Grossman, A., Davles, J., Federspiel, N., Harris, E., Hauser, C.,
Lefebvre, P., McDermott, J.P., Shrager, J., Sliflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardii Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
Parchia Plants; project phase 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   69 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys------GlyProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            231 ACAAAGTTACATACACGG -----AAGGACCAGCGCGCTTGGCCACCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ===:::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Chlamydomonas reinhardtii"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /mol_type="mRNA"
/strain="CC-1690 wild type mt+ 21gr"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           701
32
14
45
22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
```

```
ò
                                                                                   g
                                                                                                                                                   ò
                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                  ô
                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                      BF182274 11-OCT-2000 BF182274 846 bp mRNA linear EST 31-OCT-2000 601804028F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035102 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Contact: Robert Strausberg, Ph.D.
Email: ggapbs-rémail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: M.C.E. Consortium/LLNL at:
http://image.lln.jgov
Plate: LLAM9308 row: g column: 07
High quality sequence stop: 696.
| LAM9308 row: g column: 07
High quality sequence stop: 696.
| Lam9308 row: g column: 07
High quality sequence stop: 696.
| Anni Lype="mmRNA" / strain="C57BL/63" / strain="C57BL/63" / db_xrei="taxon:10090" / dcone="taxon:10090" / 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /clone_lib-"NCI_CGAP_Mam5"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_l: Sall; Site_2: Not1; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Robin Humphreys, NIH"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         709 GGTTCTTCAGTCCCAAGTGGGCTGGAGAAAGTGAACACAGAGAAAGAGGGACGGTCCCTCA 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nGlySerArg-----SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVa 120
405 GAAGGGGACGGGAATICCGTTTGCGGAAAAGTGAGCGCCCCAAGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mus musculus
Sukaryota; Metazoa; Chordata; Craniata: Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
1 (bases 1 to 846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         71 GlyAlaGlyThrArg-ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NIH MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       846
13
36
31
91
                                                             149 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 161
                                                                                               /Lissue_type="tumor, gross tissue"
/dev_stage="7 months"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     210 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 nValAspLysAspLeuValGlyTrp-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-12 (1-195) x BF182274 (1-846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241 g
                                                                                                                                                                                                                                                                                                                                                                                                                   Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                              BF182274.1 GI:11060416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  218 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33
96.00
47.20%
36.80%
9.40%
                                                                                                                                                                                                                                                                                                       mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unpublished
                                                                                                                                                                                                                                                                                                                                      BF182274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                EST
                                                                                                                                                                                                                                                                                                                                ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                              BF182274/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                à
                                                                                                                    g
```

Search completed: August 31, 2003, 04:27:24 Job time : 1896.92 secs

Protein sequence of Peptide encoded by Sequence encoded in Sequence encoded in Peptide encoded by Protein encoded by HCV-1 polyprotein.

```
901.5
901.5
901.5
901.5
901.5
901.5
901.5
901.5
901.5
                                                                                                                                                                                                                                                                                                                                            899
 Hepatitis C virus
                                                                                                                           updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       chance to have a result being printed,
                                                                                                                                                                                                                                                                                                                                                                                                        A_Geneseq_19Jun03:*

| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
| SIDSI/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1983.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1986.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1986.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1980.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqgg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqggeneseqg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqggeneseqg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqggeneseqg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqggeneseqg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqggeneseqg-embl/AA1990.DAT:*
| SIDSI/gcgdata/geneseqggeneseqgenebbl/AA2000.DAT:*
| SIDSI/gcgdata/geneseqggeneseqggenebbl/AA2000.DAT:*
| SIDSI/gcgdata/geneseqggeneseqggenebbl/AA2000.DAT:*
| SIDSI/gcgdata/geneseqggeneseqggenebbl/AA2000.DAT:*
| SIDSI/gcgdata/geneseqgenebbl/AA2000.DAT:*
| SIDSI/gcgdata/geneseqgenebbl/AA2000.DAT:*
| SIDSI/gcgdata/geneseqgenebbl/AA2000.DAT:*
| SIDSI/gcgdata/geneseqgenebbl/AA2000.DAT:*
                                                                                          August 30, 2003, 17:42:58; Search time 44.6227 Seconds (without alignments) 700.745 Million cell updates,
                                                                                                                                                                                 ......VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to l
score greater than or equal to the score of the result bein
and is derived by analysis of the total score distribution
                                                                                                                                                                                                                                                                                     1107863
           version 5.1.6
- 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                  Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                       1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SUMMARIES
                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                      protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB15222
AAB15220
AAB15226
AAB15224
AAB15224
AAB15225
AAB15225
AAB152120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15221
                                                                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                   US-09-965-594-14
1032
1 MKKKGSVVIVGRINLSGDTA.
             GenCore
Copyright (c) 1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                              length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12222221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 197
195
195
197
197
197
195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0
                                                                                                                                                      Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                sed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1032
1015
998
993
985
973
963
963
                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
                                                                                                                                                                                                                                                                                                             Minimum DB s
Maximum DB s
                                                                      OM protein
                                                                                                                                                                                                                                                         Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                 Database
                                                                                                  ou
o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7 m 4 m 6 m 8 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
```

ALIGNMENTS

AAW77397 ABP71460 AAG99289

ABG72261 AAR25135 AAR28582 AAR08123 AAW93482

68981

AAE1

AAE19905 AAE21837 AAE21838

Composite hepatiti
Amino acid sequenc
Polyprotein encode
Compiled HCV seque
Hepatitis C virus
HCV polyprotein.
HCV-1 NS3/4a wutan
HCV-1 NS3/4a confo
HCV polypeptide 1.
HCV amino acid seq
HCV polypeptide 2.
HCV polypeptide 3.
HCV MS3 protein.
HEPATITIS C virus

AAW344B0 AAW4003B AAE22049 AAU84597

AAR66995 AAB15211

AAB18541 AAR21519

AAB15221	AAB15221
A	AAB15221 standard: protein: 197 AA.
X	
AC	AAB15221;
×	
i X	19-DEC-2000 (first entry)
DE	Hepatitis C virus NS4A-NS3 fusion protease #3.
×	
¥	Hepatitis; NS3 protease; viral replication; chronic liver disease:
X	liver failure; liver cancer; mutant; mutein.
×	
SO	Hepatitis C virus.
SO	Synthetic.
×	
PN	WO200040707-A1.
×	
PD	13-JUL-2000.
×	
ΡF	06-JAN-2000; 2000WO-DS00345.
×	
PR	08-JAN-1999; 99US-0115271.
×	
ΡA	(BRIM) BRISTOL-MYERS SQUIBB CO.
×	
ΡΙ	Wittekind M, Weinheimer S, Zhang Y, Goldfarb V;
×	
DR	WPI; 2000-465976/40.
DR	N-PSDB; AAA73330.
×	
Ē	Modified benatitie C wirns (HCU) NG3 protease compatitions at least 1

us-09-965-594-14.rag

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                     TCINGVCWTVYHGAGTRTIASPRGPVIQMYTNVDRDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                   LVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                           substitution of a hydrophobic alpha-helix \theta amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                            1 MKKKGSVVIVGRINLSGDTAYAQOTRGEEGCQETSQTGRDKNOVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protease; viral replication; chronic liver disease;
liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                  0;
                                                                                                                                                                                                                                                                                                     Length 197;
                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                    100.0%; Score 1032; DB 21; 100.0%; Pred. No. 1.1e-98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Goldfarb V;
                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus NS4A-NS3 fusion protease #4
                                                                                                                                                                                                                                                                                                                                Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15222 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                          Claim 23; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-465976/40.
N-PSDB; AAA73331.
                                                                                                                                                                                                                                                                                                                 al Similarity
197; Conserv
                                                                                                                                                                                                                                                                       197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15222;
                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                variant
                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15222
 a
                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
```

```
ö
                                                                                                            The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV an lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 TCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 TCINGVCWTVYHGAGTRIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 HKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDRNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                Score 1015; DB 21;
Pred. No. 6.2e-97;
0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15220 standard; protein; 195 AA.
                                                                                  Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                              Query Match 98.4%;
Best Local Similarity 98.5%;
Matches 194; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                                                                                 197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic.
                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15220;
                                                                                                                                                                                                                                                                                                      variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
```

~

```
a
                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
ò
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                    The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as
                                                                                                                                                                                                                                                                                                                                                                                             121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                           TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                  1 MKKKGSVVIVGRIVLNG--AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 58
         Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                        well as enabling structural studies of the protease and
protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                     1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                               59 TCINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication: chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                        Score 998; DB 21; Length 195; Pred, No. 3.5e-95;
                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15226 standard; protein; 197
                                                                Claim 23; Fig 12; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                        96.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                               Matches 193; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-465976/40
                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                   195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   38-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                   119
                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15226;
                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                               variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix {\bf 0} amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96.2%; Score 993; DB 21;
97.5%; Pred. No. 1.2e-94;
clve 0; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        liver failure; liver cancer; mutant; mutein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ź
                                                                                                                                                                                   Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M, Weinhelmer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       st Local Similarity 97.5 tches 192; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
N-PSDB; AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         W0200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15223;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15223
```

```
AAB15225;
                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                            variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
   a
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (RCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                    61 TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                 121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 60
                                                                                                                                                                                                                                                                                                                                                             1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA 60
                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                  0
                                                                                                                                                                                                                                                                                          Score 985; DB 21; Length 197;
Pred. No. 8e-94;
0; Mismatches 6; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15224 standard; protein; 197 AA
                                                                                            Claim 23; Fig 15; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                          95.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                 191; Conservative
WPI; 2000-465976/40.
N-PSDB; AAA73332.
                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                      197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15224;
                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                variant
                                                                                                                                                                                                                                                                                                       Local
                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                    å
                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
ö
                                                                                                                                                                                                                                                                       The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NS4A well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 TCINGVCWTVYHGAGTRTIASPRGPVIQMYTNVDRDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 973; DB 21;
Pred. No. 1.4e-92;
2; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15225 standard; protein; 197 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                           Claim 23; Fig 16; 66pp; English,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94.38;
95.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 95.4'
Matches 188; Conservative
2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                          N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
```

S

```
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29-JAN-1993;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAR-2003
27-JAN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     W09315193-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-JAN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAR40120;
                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR40120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
 ö
                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                  The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins; its useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                              TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                          LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                                                                                                                                                                                                                                       9
                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                            1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                93.3%; Score 963; DB 21; Length 197; 94.9%; Pred. No. 1.5e-91; Live 2; Mismatches 8; Indels (
Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15212 standard; protein; 195 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                        Matches 187; Conservative
                      WPI; 2000-465976/40.
N-PSDB; AAA73334.
                                                                                                                                                                                                                                                                                                                             Sest Local Similarity
                                                                                                                                                                                                                                                                                          197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15212;
                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                     variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB1521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
61 TCINGVCWTVTHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                           Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of MS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 HKKGSVVIVGRIVLNG--AYAQQTRGLLGCIITSLIGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV genomic amino acid sequence isolated from infected human LG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C Virus; Non-A, non-B hepatitis Virus; HCV; NANBHV; human growth hormone; HGH; secretion signal; fusion protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 195;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6; Indels
   Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   / Match 92.9%; Score 959; DB 21;
Local Similarity 95.4%; Pred. No. 3.9e-91;
nes 188; Conservative 1; Mismatches 6;
   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAR40120 standard; Protein; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                              Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AVDFIPVESLETTMRSP 197
Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           93WO-DS00907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     92US-0830024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (updated)
(first entry)
                                                                          WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C Virus
                                                                                                             N-PSDB; AAA73328
```

```
THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ď
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones 14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1065 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDRDLVGWPAPQGARSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                      RNA was isolated from the plasma of a HCV seropositive human (designated "LG") and cDNA was prepared from it. The cDNA was PCR amplified using specific primers with sequences based on the prototype HCV-1 CDNA sequence (GENBANK M62121). Further amplification using nested primers resulted in 7 adjacent HCV DNA fragments which could be assembled into a full-length sequence. The DNA sequence was determined and translated into the genomic amino acid sequence. Comparison of the LG genomic amino acid sequence with that from HCV-1 showed 134 amino acid differences.
                                                                                                                                                                                                        New plasmid pHCV-162 is a mammalian expression systems for HCVD proteins - useful for diagnosing HCV infection and as vaccines for preventing HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87.5%; Score 902.5; DB 14; Length 3011;
85.8%; Pred. No. 1.1e-83;
Live 9; Mismatches 11; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                               Frail DE;
                                                               Devare SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 1; Page 39-49; 100pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTRCVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP92041 standard; protein; 1766 AA
                                                               Desai SM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87US-0122714.
87US-0139886.
88US-0161072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 85.8 tes 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAP92041;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAP92041

XXX

XXX

AC AAP

XXX AAP

XXX BG

XXX

XXX

XXX Hep

XX
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
```

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                     It is the sequence encoded in the open reading frame of hepatitis C virus cDNA inserts in clones 141,m 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 31s, 32b, 25c, 14c, 8f, a33f, 33g and 39c. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                       Claim 13; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 10; Length 1766;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Protein sequence of hepatitis c virus composite cDNA
                                                                                                                                                                                                                                  - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87.4%; Score 901.5; DB 1
85.8%; Pred. No. 6.9e-84;
                                                                                                                                                                                                                                                                                                                                                                                                                     (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP90158 standard; protein; 1786
                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88GB-0027024
                                                                                                                    Kuo
                                                                                                                                                                                                                Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus; vaccine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (updated)
                                                                                                                Houghton M, Choo QL,
                                                                                                                                                      WPI; 1989-159274/22.
                                                                              (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1766 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                          N-PSDB; AAN92097
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAR-2003
10-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP90158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sest Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
```

~

```
WPI; 1989-159274/22.
N-PSDB; AAN92103.
WPI; 1989-215054/30.
N-PSDB; AAN90331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAP92047;
                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DAX PERMENTAL DESCRIPTION OF SECULAR S
                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ::| :::| :::| :::| 289 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 AAQIFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53
                                                                                                                                                                                                                                                              The sequence is encoded by the composite cDNA of AAN90327. These antigens react with antibodies in patients with non-A non-B hepatitis (NANBH). They can be used to diagnose HCV-induced NANBH, to raise antibodies for immunoasasy or treatment, or to produce vaccines. (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                    prodn. of polynucleotide probes, diagnosis, prevention and treatment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KRGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                         9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                    DB 10; Length 1786;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus; clone 12f; clone 15e; probe; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Peptide encoded by composite hepatitis C virus cDNA.
                                                                                                                                                                                                                                                                                                                                                                                                      Score 901.5; DB 10;
Pred. No. 7e-84;
                                                                                                                                                                                                                                                                                                                                                                                                      87.4%; Scc. No. , c
85.8%; Pred. No. , c
9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP90164 standard; protein; 2261 AA
                                                                                                                                                  Hepatitis C virus gene - used for polypeptide(s) and antibodies for of infection.
                                                                                                                                                                                                                          Disclosure; fig 26; 30pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Houghton M, Choo QL, Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
                                                          Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (CHIR ) CHIRON CORPORATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (updated)
(first entry)
               (CHIR ) CHIRON CORPORATION
                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match 87.4
Best Local Similarity 85.8
Matches 175; Conservative
                                                          Choo QL,
                                                                                           WPI; 1989-215054/30
N-PSDB; AAN90327.
                                                                                                                                                                                                                                                                                                                                                                                 Sequence 1786 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    30-DEC-1987;
26-FEB-1988;
26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-MAR-2003
01-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GB2212511-A
                                                        Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26-JUL-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAP90164;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
AAP90164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Š
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                           439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones
12f through 15e.
                                                                                                                                                                                                                                                                                                                                                                                                                                          ::| :::| : | :::| 380 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                              The sequence is the peptide encoded by the composite hepatitis C virus (HCV) CDNA of AAN90331. The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, our to produce vaccines.
(Updated on 25-WAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                   9; Gaps
Hepatitis C virus gene – used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and
                                                                                                                                                                                                                                                                                                                   Score 901.5; DB 10; Length 2261;
Pred. No. 9.6e-84;
9; Mismatches 11; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAP92047 standard; protein; 2301 AA
                                                                                        Disclosure; fig 32; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                      87.48;
85.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                         st Local Similarity 85.8 tches 175; Conservative
                                             treatment of infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                           2261 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
```

6

Best Loca Matches

셤 ò 셤 õ g ó 셤

```
It is the sequence encoded in the open reading frame of hepatitis C virus (HCV) cDMA inserts in clones K9-1 through 15s. It is antiqunic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, (Updated on 25-MAR-2003 to correct PR field.)
(Updated on 25-WAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                           54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                         555 REGREILLGPADGMVSKGWRLLAPITATAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                                                                                                                                                                                                Length 2436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus; clone 15e; clone k9-1; probe; vaccine.
                                                                                                                                                                                                                                                                             11; Indels
                                                                                                                                                                                                                                            DB 10;
                 and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                            87.4%; Score 901.5; DB 1085.8%; Pred. No. 1.1e-83; tive 9; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Peptide encoded by composite hepatitis C cDNA.
                                                 Claim 13; Figure 47-1 - 47-8; 139 pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP90288 standard; protein; 2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87US-0122714.
87US-0139886.
88US-0161072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
Purified hepatitis C virus
                                                                                                                                                                                                                                                          Best Local Similarity 85.8
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (updated)
(updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              choo OL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1989-215054/30.
                                                                                                                                                                                                               2436 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAN90336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003
19-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GB2212511-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-OCT-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26-JUL-1989
                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAP90288;
                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                            380 RRGREILLGPADGMVSKGWRLLAPITAYAQQIRGLLGCIITSLIGRDKNQVEGEVQIVST 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        It is the sequence encoded in the open reading frame of hepatitis C virus (MCV) cDNA inserts in clones 12f through 15e. It is antigenic and could be used in immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A,
                                                                                                                                                                                                                                                                                                                                                                                          AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                           114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence encoded in the hepatitis C virus (HCV) cDNA inserts in clones K9-1 through 15e.
                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN------LSGDTAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                            Length 2301;
                                                                                                                                                                                                                                                                                              6
                                                                                                                                                                                                                                                            87.4%; Score 901.5; DB 10; Length
85.8%; Pred. No. 9.9c-84;
tive 9; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus (HCV); non-A, non-B hepatitis (HANBH)
                                   - and associated nucleic acids and polypeptide(s)
                                                                                                                                                              non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
(Updated on 25-MAR-2003 to correct PI field.)
                                                                  Claim 13; Figure 32-1 - 32-7; 139 pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CIRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAP92050 standard; protein; 2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               87US-0122714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88US-0271450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                  Purified hepatitis C virus
                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (updated)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Houghton M, Choo QL,
                                                                                                                                                                                                                                                                           Local Similarity
nes 175; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1989-159274/22.
N-PSDB; AAN92106.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (CHIR ) CHIRON CORP
                                                                                                                                                                                                                               2301 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .8-NOV-1987;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-MAY-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26-0CT-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30-DEC-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EP318216-A.
                                                                                                                                                                                                                                                                                                                                                                                          54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP92050;
                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                               Query Match
```

RESULT 14

AAP92050
XX
AC
AAP
AAP
XX
AAP
DT
CS-1
DT
CS-2
DT
CS-3
DT
CS-3
DT
CS-4
DT
CS-6
DT
CS-6
DT
CS-6
DT
CS-6
DT
CS-6
DT
CS-6
DT
CS-7

```
1;
                                                                                                                                                                                                                               54 AAQIFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                  The sequence is the peptide encoded by the composite hepatitis C virus (HCV) cDNA of AAN90336. The polypeptides are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines.

(N.B. This record was resulmnitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                     9; Gaps
                                                                                                                                                                     Query Match

87.4%; Score 901.5; DB 10; Length 2436;
Best Local Similarity 85.8%; Pred. No. 1.1e-83;
Matches 175; Conservative 9; Mismatches 11; Indels 9;
            Disclosure; fig 47-1 to 47-8; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                            174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                             2436 AA;
                                                                                                                                              Sequence
င်
                                                                                                                                                                                                                                                          QQ
                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

Search completed: August 30, 2003, 19:12:23 Job time: 45.6227 secs

```
3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                175; Conservative
73.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
õ
                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein
genome polyprotein
genome polyprotein
genome polyprotein
genome polyprotein
genome polyprotein
polyprotein - dour
polyprotein - marm
hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                       polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                           angio-associated m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  nitrate/nitrite se
probable aromatic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  beta transducin ho
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         transferrin-bindin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hypothetical prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hypothetical prote
probable exported
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hypothetical prote
ferredoxin-nitrite
                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nitrate/nitrite se
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  probable periplasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyketide synthas
                                                              (without alignments)
1168.492 Million cell updates/sec
                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.
                                                                                                  1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                     August 30, 2003, 19:02:22 ; Search time 16.2134 Seconds
                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                       genome
                                                                                                                                                                                                                                                                                                                                                 genome
                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                       genome
      GenCore version 5.1.6
(c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                        hits satisfying chosen parameters:
                                                                                                                                         283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                             Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                      - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                           T08839
B71360
I39383
E75392
B81104
C81911
                                                                                                                                                                                                                                                                                                                                                GNWVC3
S40770
GNWVCH
                                                                                                                                                                                                                                                                                                                                                                                                   $18030
JC5620
JQ1303
GNWVJ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F81196
B71284
S28261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       T17410
T06118
AH1030
C72705
S38789
                                                                                                                  BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                    08841
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 r49806
                                                                                                                                                                        DB seq length: 0
DB seq length: 200000000
                                                                                    US-09-965-594-14
                                                                                                                                                                                                                                                                                                                                  图
                                                                                                                                                                                                                                                                                                                  %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                               3011
30111
30111
30100
30100
30100
30100
4950
4950
5900
5900
5900
5900
               Copyright
                                                                                                                                                                                                                                    piri:*
pir2:*
pir3:*
                                                                                                                                                                                                                              PIR_76:*
                                                                                                                                                                                                                                                                                                                                                                Fotal number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               76.5
                                                                                                                                                                                                                                                                                                                                  Score
                                                                                                                                                                                                                                                                                                                                               80.5
80.5
78.5
78.5
77.5
                                                                                            Perfect score:
                                                                                                                   Scoring table:
                                       OM protein
                                                                                                    Sequence:
                                                                                                                                          Searched:
                                                                                                                                                                                                                              Database
                                                                                                                                                                         Minimum
                                                                                                                                                                                 Maximum
                                                                                                                                                                                                                                                                                                                                                                               9876543210987654321098765
                                                      Run on:
                                                                                                                                                                                                                                                                                                                                 ĕ.
                                                                                                                                                                                                                                                                                                                         Result
```

```
A.Accession: A39166
A.Molecule type: mRNA
A.Rocassiones: 1-3011 < CHO>
A.Rocassiones: 1-3011 < CHO
A.Cossiones: 1-3011 < Choose B. W. B. Poutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 1131-1141, 1992
A.Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A.Reference number: PQ0393; MUID:92268871; PMID:1316939
A.Rocassion: PQ0403
A.Rocassion: PQ0403
A.Rocassion: PQ0403
A.Rocassion: DQ0403
A.Rocassion: SQ10128
A.Rocassion: SQ10128
A.Rocassiones: 1577-1633 < CHA>
A.Cossiones: DBBJ-101128
A.Rocassiones: DBBJ-101128
                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (strain HCV-1)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                        C;Species: hepatitis C virus
C;Date: 30-Sep-1992 *sequence_revision 30-Sep-1992 *text_change 19-Jan-2001
C;Accession: A39166; PQ0403; PQ0404
R;Choo; Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.;
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A;Title: Genetic organization and diversity of the hepatitis C virus.
A;Reference number: A39166; MUID:91172826; PMID:1848704
                                                                                                                             zona pellucida gly
hypothetical prote
hypothetical prote
hypothetical prote
ferredoxin nitrite
HIV-1 retropepsin
                                                             glycosyl transfera
env polyprotein -
transferrin-bindin
                                probable regulator hypothetical prote
 transferrin-bindin
                                                                                                              orsellinic acid sy
                                                                                                                                                                                                                                               hypothetical prote
                                                                                                                                                                                                                              conserved hypothet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87.4%; Score 901.5; DB 1; Length 3011; 85.8%; Pred. No. 3.7e-75;
                                                                                                                                                                                                                                                                                               ALIGNMENTS
               S61441
C96033
D876033
D876033
C81832
T38481
S70401
S70401
AH2406
S30922
G8NLGL
G8NLGL
911
329
329
1334
1334
603
603
415
415
424
755
846
1165
317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
```

Gaps 6

11; Indels

9; Mismatches

```
A Residues: 1-3011 <a href="https://doi.org/10.10/">https://doi.org/10.10/</a>
A Reference: 08:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329738
R) Inchauspe, G: Zebedee, S: Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
R: Inchauspe, G: Zebedee, S: Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
R: Reference number: A41546; MUID:g202-10296, 1991
A; Reference number: A41546; MUID:g202256; PMID:1658800
A; Contents: annotation
A; Reference number: A41546; MUID:g202255; PMID:1658800
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: ATP; capsid protein; executed <a href="http://doi.org/">http://doi.org/</a>
C; Reywords: ATP; capsid protein; executed <a href="http://doi.org/">http://doi.org///doi.org///doi.org///doi.org///doi.org///doi.org/</a>
F; 102-389/Product: envelope protein M *status predicted <NS1>F; 1007-161s/Product: monstructural protein NS1 *status predicted <NS2>F; 1007-161s/Product: nonstructural protein NS1 *status predicted <NS2>F; 1007-161s/Product: nonstructural protein NS2 *status predicted <NS2>F; 1312-1317/Region: nucleotide-binding motif A (P-10op)
F; 1312-1317/Region: nucle
                                                                                                                     genome polyprotein - hepatitis C virus (strain H)
N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4: nonstructural protein NS4b; nonstructural protein NS5
C; Species: hepatitis C virus
A; Note: host Homo sapiens (man)
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: A36814; A41546
S; Inchauspe, G; Zebedee, S; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
S; Date: 31-Dec-1992 #sequence contains to the human prototype strain H of hepatitis C virus
A; Psecrence number: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GNWVIW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A. Experimental source: 180 at H-J1

C. Superfamily: hepatitis C virus genome polyprotein

C. Superfamily: hepatitis C virus genome polyprotein

C. Superfamily: hepatitis C virus genome polyprotein

F. Steywords: ATP: glycoprotein: hydrolase; nucleotide CPC>

F. 115-Product: capsid protein C *status predicted <CPC>

F. 116-191/Product: capsid protein M *status predicted <EPM>
F. 116-191/Product: major envelope protein B *status predicted <MEE>
F. 192-389/Froduct: major envelope protein NSI *status predicted <NSI>
F. 1007-1615/Product: nonstructural protein NSI *status predicted <NSI>
F. 1007-1615/Product: hepacivirin *status predicted <NSI>
F. 1313-1317/Region: nucleotide-binding motif A (P-loop)

F. 1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Cross-references: EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221587
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Isuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The S'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protein M; hepacivirin (EC 3.4.21.98) (nonstructu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           113
1005 RRGREILLGPADGMYSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYIKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                               AAOTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C;Species: hepatitis c virus
C;Species: hepatitis c virus
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C;Accession: S40770; PC1285
R;Okamoto, H.
submitted to the EMBL Data Library, March 1992
A;Reference number: S40770
A;Reference number: S40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Molecule type: genomic RNA
A;Residues: 1-513 <OK2>
A;Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 .;
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87.3%; Score 900.5; DB 1;
85.8%; Pred. No. 4.5e-75;
ative 8; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 CIRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                     174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: genomic RNA A; Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mar
Local S...
175;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                  g
                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
```

```
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruu protein NS4s; nonstructural protein NS5 c; Species: hepatitis C virus
A;Note: hepatitis C virus
A;Note: hepatitis C virus
C;Date: hepatitis C virus
A;Note: hepatitis C virus
A;Note: hepatitis C virus
C;Date: A virus
C;Date: J12c-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C;Accession: A40244
R;Chen, P.J; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                               54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                          Gaps
                                                       ..
Length 3011;
                                                       Indels
85.9%; Score 886.5; DB 1; 83.8%; Pred. No. 9.2e-74;
                                                       Conservative 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 CTRGVAKAVDFIPVESLETTMRSP 197
```

```
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 AAQTFLATCINGVCWTVYRGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                               21;
                                                                             82.4%;
ilarity 77.5%;
Conservative 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  156; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                 Query Match
Best Local Similarity
Matches 158; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                     54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GNWVCJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein - hepatitis C virus

N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; limar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
C; Accession. 38466; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; J.; Vitol. 65, 1105-1113, 1991
A; Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; J.; Vitol. 65, 1105-1113, 1991
A; Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; J.; Vitol. 65, 1105-1113, 1991
A; Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; J.; Vitol. 65, 1105-1113, 1991
A; Takamizawa, A.; Mori, C.; Puke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Product form human A; Reference number: A38465; MUID:91140698; PMID:1847440
A; Accession: A38465; MUID:91140698; PMID:1847440
A; Residues: I. 3010 < TAKA

                                                                                                                                                                                             Across-references: 03:M84754
C:Superfamily: hepatitis C virus genome polyprotein glycoprotein; hydrolase; nonstructura C:Superfamily: hepatitis C virus genome polyprotein glycoprotein; hydrolase; nonstructura C:Superfamily: hepatitis C virus genome protein; glycoprotein; hydrolase; nonstructura F;116-191/Product: capsid protein C *status predicted <CPC>
F;116-191/Product: envelope protein M *status predicted <MEE>
F;390-729/Product: major envelope protein E *status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 *status predicted <NS1>
F;1007-1615/Product: nonstructural protein NS2 *status predicted <NS2>
F;1007-1615/Product: hydrolase protein NS2 *status predicted <NS2>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1616-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A:Title: The Talwanese hepatitis C virus genome: sequence determination and mapping the A:Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82.8%; Score 854.5; DB 1; Length 78.9%; Pred. No. 8.7e-71; tive 18; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match 82.8%; Score 854.5;
Best Local Similarity 78.9%; Pred. No. 8.7e
Matches 161; Conservative 18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1316-1319/Region: DEXH motif
                                                                                                                             A; Molecule type: genomic RNA A; Residues: 1-3010 <CHE>
                                                                                              A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
```

```
A Residues: 2650-2707 (ARA)

A Reperfamily: a cleavage sites of this polyprotein have not been determined.

C Comment: The cleavage sites of this polyprotein

C Superfamily: hepatitis C virus genome polyprotein

F 2-115/Product: capsid protein & *status predicted (ARE)

F 116-191/Product: envelope protein & *status predicted (ARE)

F 130-1006/Product: nonstructural protein NSI *status predicted (NSI)

F 130-1006/Product: nonstructural protein NSI *status predicted (NSI)

F 1230-1237/Region: nucleotide-binding motif & (P-loop)

F 1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                R;Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Suglmura, T.; Sh
Proc. Natl. Aadd. Sci. U.S.A. 87, 9524-9528, 1990
A;Title: Wolecular cloning of the human hepatitis C virus genome from Japanese patie
A;Reference number: A39253; MUID:91088550; PMID:2175903
A;Accession: A39253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F:1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F:205,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule type: genomic RNA
A;Residues: 1-3010 <KAT>
A;Cross-references: 08:090208; NID:g221610; PIDN:BAA14233.1; PID:g221611
A;Cross-references: 08:090208; NID:g221610; PIDN:BAA14233.1; PID:g221611
B;Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence A;Reference number: PS0085
A;Reference number: PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                  1125 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPFGHAVGIFRAAV 1184
                                                                                                                                                                                                                            114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein E;
NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E;
protein NS4a: nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C;Accession: A39253; PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KRGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKWQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                     AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                               3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSI
                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
Score 850.5; DB 1;
Pred. No. 2.1e-70;
1; Mismatches 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            82.4%; Score 850.5; DB 176.5%; Pred. No. 2.1e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 CTRGVAKAVDFIPVESLETIMRSP 197
```

```
A; Molecule type: genomic RNA
A; Residues: 1-547, "7',549-621, "V',623-624, "S',626-652, "DL',655-761, "T',763-782 < HOW>
A; Residues: 1-547, "7',549-621, "V',623-624, "S',626-652, "DL',655-761, "T',763-782 < HOW>
A; Residues: 1-547, "7',549-621, "V',623-624, "S',626-652, "DL',655-761, "T',763-782 < HOWS
A; Cross-references: Explaint With the nucleotide translation
A; Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 3
as Trp, and TTC for residue 71 as Ser
A; Note: sequence extracted from NCB1 backbone (NCBIN:121747, NCBIP:121748)
C; Superfamily: hepatitis C virus genome polyprotein
F; 115-115/Product: capsid protein C * *status predicted < CPC>
F; 116-191/Product: capsid protein E * *status predicted < NS1>
F; 130-1006/Product: nonstructural protein NS2 * *status predicted < NS2>
F; 130-1006/Product: nonstructural protein NS2 * *status predicted < NS2>
F; 130-1237/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif B
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 23-Mar-2001 C;Accession: 318030; 333570; A48332; S18029 S:33570; A48332; S18029 S:33570; A48332; S18029 S:33570; A48331; S18029 S:33570; A48331; Sission S:30570; Submitted to the EMBL Data Library, September 1991 A;Reference number: Sission S:318028 A;Reference number: S18038
                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Experimental source: isolate JK1 from an individual
R; Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A; Title: Sequence analysis of putative structural regions of hepatitis C virus isolat
A; Feference number: A48332; MUID:93119270; PMID:8380322
A; Accession: S33570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS5 *status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                    A;Wolecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross-references: EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 76.5%;
Matches 156; Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Status: preliminary
A;Status: preliminary
A;Status: preliminary
A;Status: preliminary
A;Rolecule type: DNA
A;Residues: 1-3010 crIAN>
A;Coss.references: GB:D11168; GB:D01171; NID:g221612; PIDN:BAA01943.1; PID:g221613
A;Cross.reference extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C;Superfamily: hepatitis C virus genome polyprotein bidding; P-loop; polyprotein; serin C;Superfamily: hepatitis C virus genome polyproteide bidding; P-loop; polyprotein; serin C;Superfamily: hepatitis C virus genome polyproteide bidding; P-loop; polyprotein; F;116-191/Product: envelope protein W *status predicted cREN>
F;116-191/Product: monistructural protein NS1 *status predicted cNS2>
F;130-1066/Product: nonstructural protein NS2 *status predicted cNS2>
F;120-1237/Region: nucleotide-bidding mouif A (P-loop)
F;1312-1317/Region: bEXH motif
F;1616-1862/Product: nonstructural protein NS4 *status predicted cN4A>
F;1616-1862/Product: nonstructural protein NS4 *status predicted cN4A>
F;1863-2013/Product: nonstructural protein NS4 *status predicted cN4S>
F;1863-2013/Product: nonstructural protein NS5 *status predicted cN85>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (strain JT)
N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: A;Reference number: A45573; MoID:92295714; PMID:1318627
A;Accession: A45573
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1065 ATGSFLATCVNGVCWTVZHGAGSKTLAGPKGPITQMYINVDQDLVGWPAPPGARSWTPCT 1124
                                                                                                                                                                            54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                   CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Species: hepatitis C virus
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C;Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tch 81.8%; Score 844.5; DB 1 al Similarity 77.5%; Pred. No. 7.4e-70; 158; Conservative 20; Mismatches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CTRCVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                      174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 158; Conserva
                                                                                                                                   114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 8
S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                   ð
                                                                                                                                                                                                                 g
```

se

```
Richamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Blochem, Blochbys, Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A;Reference number: JC5620; MUID:97366593; PMID:9223423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                         1124
                                                                                                                                                                                                           1064
                                                                                                                                                                                                                                                                                                                                                                                                                                        1125 YGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C.Species: hepatitis C virus
C.Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C.Accession: JC5620
                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                   54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                             Gaps
80.3%; Score 828.5; DB 1; Length 3010;
76.5%; Pred. No. 2.3e-68;
Live 20; Mismatches 19; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (isolate EUH1480)
                                                                  20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 CTRGVAKAVDFIPVESLETIMRSP 197
```

```
Matches 123; Conservative
                                                                                                       Matches 122; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F;1320-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: mRNA
A; Residues: 2678-2729 <KAT>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                           Query Match
Best Local Similarity
                                                                                                                                                             ò
                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A. Wote: the translation of the nuclectide sequence in this paper.
C. Superfamily: hepatitis C virus genome polyprotein
F; 2.115/Product: capsid protein C %status predicted <CPC>
F; 116-191/Product: envelope protein E %status predicted <MEE>
F; 384-408/Region: hypervariable %status predicted <NS1>
F; 390-730/Product: nonstructural protein NS1 %status predicted <NS2>
F; 1000/Product: nonstructural protein NS2 %status predicted <NS2>
F; 1008-1616/Product: hepacivirin %status predicted <NS3>
F; 1013-1318/Region: nucleotide-binding motif A (P-loop)
F; 1131-1130/Region: nucleotide-binding motif A (P-loop)
F; 1131-1130/Region: DEXH motif
F; 1151-1185/Product: nonstructural protein NS4 %status predicted <N4A>
F; 1864-2014/Product: nonstructural protein NS5 %status predicted <NS5>
F; 2015-3014/Product: nonstructural protein NS5 %status predicted <NS5>
F; 2015-3014/Product: nonstructural protein NS5 %status predicted <NS5>
F; 2015-3014/Product: nonstructural protein NS5 %status predicted <NS5>
F; 2016-2019/Region: interferon sensitivity determining %status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C:Species: hepatitis C virus
C:Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>F;1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
A;Cross-references: GB:Y13184
A;Experimental source: genotype 5a, which predominates in South Africa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1; Length 3014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 74.3%; Score 766.5; DB 1
Best Local Similarity 69.6%; Pred. No. 1.3e-62;
Matches 142; Conservative 25; Mismatches 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CIRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

```
Cisosa-references: GB:D10562, GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
C;Superfamily: hepatitis C virus genome polyprotein
C;Superfamily: hydrolase; nonstruct
F;1-115/Product: capsid protein C *status predicted <CPC>
F;116-191/Product: envelope protein M *status predicted <NED>
F;130-733/Product: nonstructural protein NSI *status predicted <NSI>F;34-1010/Product: nonstructural protein NSI *status predicted <NSI>F;111-11519/Product: hepativitin *status predicted <NSI>F;1214-1241/Region: nucleotide-binding motif A (P-loop)
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NO5>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2036
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA
A; Residues: 1.3033 < CMA>
A; Residues: 1.3033 < CMA>
A; Cross-references: GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; PID:g221609
A; Chan, S; W.; McCmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J.; Gen. Virol. 73, 1131-1141, 1992
A; Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A; Reference number: PQ0393; MUID:92268871; PMID:136939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         T.; Mori, S.; Hijikata, M.; Shimot
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A40250; P00359; P00559
A;Title: Pull-length sequence of a hepatitis C virus genome having poor homology to A;Reference number: A40250; MUD::3230232; PMID::1314459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                            1034 TAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGTTISGVLWTVYHGAGNKT 1093
                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               79 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                       19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65.1%; Score 672; DB 1; Length 3033; 68.7%; Pred. No. 8.4e-54; ive 24; Mismatches 32; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
A; Residues: 2678-2754 <CHA>
A; Cross-references: DDBJ:D10134
A; Cross-references: DDBJ:D10134
A; Experimental source: isolate E-bl2
B; Rato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; MC
B; Chem. Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720309
A; Accession: PQ0559
                                                                                                                                 65.3%; Score 674; DB 1; 68.2%; Pred. No. 5.5e-54;
                                                                                                                                                                                                                       26; Mismatches
```

Sun Aug

qq οy 셤 ö g 12

```
anglo-associated migratory cell protein - human c.Species: Homo sapiens (man) c.Species: Homo sapiens (man) c.Species: Homo sapiens (man) c.Species: Homo sapiens (man) c.Species: No. Sep-1996 #text_change 21-Jul-2000 c.Species: No. S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A:Residues: 1-495 <COL>
A:Cross-references: GB:AE001199; GB:AE000520; NID:g3322402; PIDN:AAC65137.1; PID:g332
A:Experimental source: strain Nichols
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pypothetical protein TP0136 - syphilis spirochete
Cippecies: Treponema palidum subsp. palidum (syphilis spirochete)
Cipace 24 Jul. 1998 #sequence_revision 24 Jul. 1998 #text_change 05-Nov-1999
Cipacession: B71360
Cipacession: Cip
                                                                                                                                                             1066 KCQPQGVWVI-----RND--GALCHGTLGRTVELDLPAELCDFRGSSGSPILCD 1112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVMSIATDGNTFVLACVP 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 -NGVCMTVYHGAG----TRIIASPKGPVIQMYINVDKDLVG-----WPAPQGSRSLTPCI 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               119 GTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT----LVGGTSKPFWLVPGGTGNNGNCG 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 C-----GSSDLFLVTRHADVIP------VRRRGDSRGSLLSPRPISYLK------ 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        175 CGGGGGGSSSSSSCIHIMLVPGGTGNNGNCGCGGGGGSSSSSSSCIHIKVENTDEQFL 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: B71360
A; Status: preliminary; nucleic acid sequence not shown; translation not shown
A; Wolecule type: DNA
                                                                              53 TAAQIFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16 SGDTAYA------QQTRGEEGCQETSQ----TGRDKNQVEGEVQIVSTAAQTFLATCI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Cross-references: GB: M95627; NID: 9870802; PIDN: AAA68889.1; PID: 9870803
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8.3%; Score 85.5; DB 2; Length 495; ilarity 23.4%; Pred. No. 3.7; Conservative 17; Mismatches 76; Indels 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       162 AGHAVGIFRAAVCTRG-----VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         152 ------GSSGGPLLCPAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGG 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-452 <RES>
                                                                                                                                                                                                                                                                                                                                                        113 TCGSSDLYLVTRHADVÍPVRRRGDSRGSLLS---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 50; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: I39383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C;Genetics:
A;Gene: TP0136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     polyprotein - marmoset hepatitis GB virus A
C;Species: 13-01-1999 *sequence_revision 23-01-1999 *text_change 17-Nov-2000
C;Accession: T08839
B;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.
A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.
A;Reference number: 216486; MUID:98120818; PMID:9460920
A;Reference number: 216486; MUID:98120818; PMID:9460920
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: genomic RNA
A;Residues: 1-2970 cRRK
A;Residues: 1-2970 cRRK
A;Residues: 1-2970 cRRK
C;Superfamally: hepatitis C virus genome Polyprotein
C;Reywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein - douroucouli hepatitis GB virus A C; Species: douroucouli hepatitis GB virus A C; Species: douroucouli hepatitis GB virus A C; Date: 20.5ep-1999 #sequence_revision 20.5ep-1999 #text_change 17-Nov-2000 C; Accession: T08841 R; Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. Virol. 79, 41-45, 1998 A; Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A; Reference number: 216486; MUID: 98120818; PMID: 9460920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3:1 | 1:1 | 1:1 | 1:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 3:1 | 
                                                                              'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              995 KISMLGRDEREHEGSIVVLGTSTTRSMGTCVNGVMYTTFHGSNARTLAGPVGPVNCRWWS 1054
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1055 PSDDVAVYPLPSGASCLEPCXCGTQSVWCIRN--DGALCHGRLSKLVELDLPTEISDFRG 1112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A:Molecule type: mRNA
A:Residues: 1-3005 <ERK>
A:Cross-references: EMBL:AF023425; NID:g2828599; PIDN:AAC40502.1; PID:g2828600
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: polyprotein
                                                                                                                                                                                                                                                                         79 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        93 VDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKG 152
TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             33 ETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTN 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVS 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 3005;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 24.9%; Score 257; DB 2; Length 300
Best Local Similarity 34.1%; Pred. No. 3.4e-15;
Matches 56; Conservative 29; Mismatches 69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SSGGPLLCPAGHAVGIFRAAVCTRGV-----AKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A:Status: translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 27.8%
Matches 62; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1113
13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139
```

ð 셤 õ g ò g

셤 ð

υŒ

Search completed: August 30, 2003, 19:20:27 Job time : 17.2134 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 9.75674 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

Title: Perfect score: Sequence:

US-09-965-594-14 1032 I MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTWRSP 197

BLOSGM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues

Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

STIMMARTES

		æ			SUMMARIES		
Result	Score	Query Match	Query Match Length	DB	ID	Description	
1	901.5	87.4	3011	-	POLG HCV1	P26664 h genome po	
7	886.5	ı,	3011	-	POLG_HCVH	h genome	
٣	854.5	N	3010	-	POLG_HCVTW	h genome	. ~
4	850.5	82.4	3010	-	POLG_HCVBK	denome	_
2	850.5	ď	3010	-	POLG_HCVJA	h genome	^
9	844.5	ä	3010	٦	POLG_HCVJT	h genome	_
7	674	65.3	3033	٦	POLG_HCVJ6	h genome	_
00	672	ď.	3033	٦	POLG_HCVJ8	h genome	_
6	85.5		485	-1	Y136_TREPA	treponema	_
10	85		321	Н	HHOA_ARATH	arabi	"
11	82		452	~	AAMP_HUMAN	013685 homo sapien	_
12	80.5		437	٦	DEGLARATH	arabi	
13	78.5	7.6	209	-	PAAD_PSEAE		,,
14	77.5		2663	-	CENE_HUMAN	Q02224 homo sapien	~
15	16		911	-	TB11_NEIMB	neis	_
	75.5	7.3	786	۲-	SN1L_HUMAN	P57059 homo sapien	_
	75		603	۳-	ENV_RSVP	rous	_
18			1705	-	PTPO_MOUSE		_
19	4		415	~	ZP3_RABIT	-	, ,
20	74.5		116	7	HYPF_AZOVI		٤.
21	7		1165	-	POL_GALV	gibb	
22	73.5	٠	263	Н	GRAK_MOUSE	035205 mus musculu	
23	۳.	٠	199	_			_
24	73	٠	253	_	CAC3_BOVIN	P05805 bos taurus	
25	73	٠	259	-	IBP1_HUMAN	рошо	_
56	72.5	•	257	-	GRAM_HUMAN	рошо	_
27	72.5		206	-	TRFE_HORSE	ednns	
28	72		629	-	VST2_HEVME	Q03500 hepatitis e	ď
29	72	•	1527	-	CA1H_MOUSE	_	_
30	71.5		248	-		mus m	_
31	71.5	6.9	248	-	TRY1_CHICK	Q90627 gallus gall	,
32	71.5	ø.	408	-1		P80146 thermus sp.	
	71	•	336	-	UL16_EBV	epstein	

				P33426 hepatitis e P19532 homo sapien	
CN4A_RAT TBP1_NEIGO	ITAl_RAT YQU3_CAEEL	CTRL_HUMAN	ENV_FLVGL	VST2_HEVPA TFE3_HUMAN	PANE_RHILO TRPD_PSEPU
844 915	1180	264	642	660	326 349
6.9 6.9	6.9 6.9	80.80	80.00	8.89	6.8 8.8
17.	17	70.5	70.5	70.5	70
3.44 3.5	36	38	40	4.2	44 45

ALIGNMENTS

```
Structure 6:89-100(1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of unwinding.".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVH
                                                                                   ò
                                                                                                                                         q
                                                                                                                                                                                                     οż
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLULAR ANINOPEPTIDASE.
CARSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/PE2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 87.4%; Score 901.5; DB 1; Length 3011;
85.8%; Pred. No. 5.5e-77;
tive 9; Mismatches 11; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                             ); IPR002518; HCV_NS2.

); IPR004109; HCV_NS3.

); IPR001409; HCV_NS4a.

); IPR001409; HCV_NS4a.

); IPR001409; HCV_NS5a.

); IPR001566; HCV_NS6a.

); IPR001566; HCV_NS6a.

); IPR0015095; RAL_POL_DS_PS.

; IPR007094; RNA_POL_DS_PS.

); IPR007094; RNA_POL_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS5s; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; VITaL_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                             HCV_env.
HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2240 2240
2364 2364
2789 2789
3011 AA; 327197 M
                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF0153;
Pfam; PF01560; HCV_NS1;
Pfam; PF01539; HCV_NS2; 1.
Pfam; PF01200; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
                             IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  175; Conserv
                                                                                                                                                                                                                                                                                                                                               Pfam; PF01543;
Pfam; PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3D-structure.
                                                                                                                       InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local
Matches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
Sort First Fritting Fritzen Frit
```

Gaps

.; 6

KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

Conservative

m

ò

```
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.99-); Protease/helicase NS3 (P70) (Hepacivitin)
(EC 3.4.99-); Protease/helicase NS3 (P70) (Hepacivitin)
NS48 (P27); Nonstructural protein NS5A (P4); Nonstructural protein
NS5B (P65) (P70) (RMA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flavivridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Rim J.L., Morgenstern R.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -I- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-I-FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-I- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1
AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -1- FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-1- FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates.";
Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-92052256; Pubmed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3011 AA.
                                                                                                                                                                                                                                                                     174 CIRGVAKAVDFIPVESLETIMRSP 197
```

```
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                              Score 886.5; DB 1;
Pred. No. 1.5e-75;
                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     85.9%;
83.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1417
1420
1436
1439
1453
1463
1480
                                                                                                                                                                                                                                                                                                                                                                                                                             1226
1233
1238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
STRAND
STRAND
HELIX
SEQUENCE
                                              CARBOHYD
CAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
TURN
STRAND
STRAND
STRAND
HELIX
TURND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
HELIX
STRAND
HELIX
TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HELIX
TURN
HELIX
TURN
              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institution. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease; 3D-structure. 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
N-LINKED (GLCNAC. ..) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENVELOPE GLYCOPROTEIN E1. ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELLULAR AMINOPEPTIDASE. CAPSID PROTEIN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS NONSTRUCTURAL PROTEIN NS NONSTRUCTURAL PROTEIN NS NONSTRUCTURAL PROTEIN NS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSFAC. 199:011.

InterPro: IPR001410: DEAD.

InterPro: IPR001552: HCV_capsid.

InterPro: IPR002521; HCV_capsid.

InterPro: IPR002519; HCV_core.

InterPro: IPR002519; HCV_core.

InterPro: IPR0002519; HCV_NS1.

InterPro: IPR0001409; HCV_NS2.

InterPro: IPR0001409; HCV_NS3.

InterPro: IPR0001409; HCV_NS4.

InterPro: IPR0001409; HCV_NS4.

InterPro: IPR0001650; HCV_NS4.

InterPro: IPR0001650; HCV_NS4.

InterPro: IPR0001650; HCV_NS4.

InterPro: IPR0001650; HCV_NS3.

InterPro: IPR0001650; HCV_NS3.

InterPro: IPR0001650; HCV_NS3.

InterPro: IPR001650; HCV_N
                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M67463; AAA45534.1; --
PIR, A36814; GRW7CH.
PDB; 1HEI; 25 NOV-98.
PDB; 1AIV; 16-FEB-99.
PDB; 1AIR; 17-JUN-98.
MEROPS; S29.001; --
MEROPS; U99.001; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    191
383
746
809
1026
1057
11057
1107
1107
1107
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
```

```
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
     SOLUTION TO THE PROPERTY OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                           01-APR'1993 (Rel. 25, Created)
01-APR'1993 (Rel. 25, Last sequence update)
01-APR'1993 (Rel. 25, Last sequence update)
01-APR'1993 (Rel. 42, Last annotation update)
Genome polyprotein (Contains capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP88) (GP70) (NS1); Protein P1 Non Nonstructural protein NS4 (P40); Nonstructural protein NS48 (P40); Nonstructural protein NS48 (P40); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P56); Nonstructural protein NS58 (P60) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                   CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            {RNA}(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LIPOPROTEIN ENVELOPE, THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
 .
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 AA
 Mismatches
                                                                                                                                                                                                                                                     EMBL; M84754; -; NOT_ANNOTATED_CDS
 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD.
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVTW
P29846;
                                                                                                                                                                                 114
                                                                                                                                                                                                                   1125
                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                 RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                              õ
                                                                                                                                              염
                                                                                                                                                                         ò
                                                                                                                                                                                                                                                       ò
```

```
ή
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS44 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CELLULAR AMINOPEPTIDASE. CORE PROTEIN (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 854.5; DB 1;
Pred. No. 1.6e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL
                              InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR00145; HCV_NS4.
InterPro; IPR00146; HCV_NS4.
InterPro; IPR001069; HCV_NS4.
InterPro; IPR0010095; RNA_POl_DS_PS.
InterPro; IPR0010095; RNA_POl_DS_PS.
InterPro; IPR0010095; RNA_POl_DS_PS.
InterPro; IPR0010095; RNA_POl_DS_PS.
InterPro; IPR0010094; RNA_POl_DS_PS.
InterPro; IPR0010095; RNA_POl_DS_PS.
InterPro; IPR001099; RNA_POl_DS_PS.
InterPro; IPR0010995; HCV_COTE, I.
Pfam; PF01560; HCV_NS1; I.
Pfam; PF01560; HCV_NS3; I.
Pfam; PF01001; HCV_NS3; I.
Pfam; PF01001; HCV_NS4; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002521;
IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3D-structure
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Simi
Matches 161;
```

us-09-965-594-14.rsp

```
1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDRNQVEGEVQVVST 1064
                                               113
                                                                                           114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                              01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 24, Last annotation update)
16-Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
16-SEP-2003 (Rel. 24, Last annotation update)
17-SEP-2003 (Rel. 24, Last annotation update)
18-SEP-2003 (Rel. 24, Last annotation protein C (Core protein C (CP68) (GP70) (Rel. 24, CP2-1);
18-SEP-2003 (Rel. 25, Protease/Helicase NS3 (P70) (Hepacivirin)
18-SEP-2003 (Rel. 25, Rel. 24, Letter and protein NS4 (P4); Nonstructural protein NS4 (P65); Nonstructural protein NS5 (P66) (P07) (Rel. 27, T. 48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIOMYTNVDKDLVGWPAPQGSRSLTPCT
                                                           X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; Pubmed-8861916;
Love R.A., Parige H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals a
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.;
"Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus (isolate BK) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691. MEDLINE-98227846; PubMed-9568891; Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M., Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.; "Complex of NS3 protease and NS4A peptide of BK strain hepatitis C virus: a 2.2-A resolution structure in a hexagonal crystal form."; Protein Sci. 7:837-847(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WEDLINE-96235224; PubMed-8647104;
BOTOWSKI P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphoxylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    trypsin-like fold and a structural zinc binding site."; cell 87:331-342(1996).
                                                                                                                                                       174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-91140698; Pubmed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          from human carriers."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                           POLG_HCVBK
                                                                                                                                                                                                                POLG_HCVBK
                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                            셤
                                                                    a
                                                                                           ŏ
                                                                                                                   a
                                                                                                                                           ò
                                                                                                                                                                 셤
                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CAPSID PROTEIN C (POTEWIAL).
MATRIX PROTEIN C (POTEWIAL).
MAJOR ENVELOPE PROTEIN E (POTEWITAL).
NONSTRUCTURAL PROTEIN NS1/F2 (POTEWITAL).
PROTEASEPHELICASE NS3 (POTEWITAL).
PROTEASEPHELICASE NS3 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWILA).
NONSTRUCTURAL PROTEIN NS4A (POTEWILA).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BY AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MENA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein, Glycoprotein; Transferase, RNA-directed RNA polymerase, Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro: IPR000445; HCV.N44.
InterPro: IPR001496; HCV.N34.
InterPro: IPR001496; HCV.N34.
InterPro: IPR001266; HCV.N34.
InterPro: IPR001266; HCV.RdRP.
InterPro: IPR001094; RNA_POL_DS_PS.
IPEAM: PF01539; HCV_COTE: 1.
IPEAM: PF01509; HCV_NS2: 1.
IPEAM: PF01500; HCV_NS3: 1.
IPEAM: PF01006; HCV_NS3: 1.
IPEAM: PF01001; HCV_NS4b: 1.
IPEAM: PF0100998; VITAL_RORP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002531; HCV_NSI.
IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                      EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               115
191
383
729
1006
1615
1862
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            369
1083
1107
1165
1237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                              1A1Q; 25-MAR-98
1JXP; 14-JAN-98
1NS3; 08-APR-99
1C2P; 15-NOV-09
1GS5; 09-APR-02
1GX5; 09-APR-02
1GX6; 10-APR-02
1GX6; 20-JUN-00
80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR00141Q;
                                                                                                                                                                                                                                                                                                                            GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            347
1083
1107
1165
1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1007
1616
1863
2014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    u39.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MD-structure.
                                                                                                                                                                                                                                                                                                                               A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                       PDB;
PDB;
PDB;
PDB;
PDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDB;
    --
```

1185 CTRGVAKAVDFVPVESMETTMRSP 1208

```
54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                       9; Gaps
                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                             (POTENTIAL). (POTENTIAL). (POTENTIAL).
                        (POTENTIAL)
                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                             Length 3010;
                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                 327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                     (GLCNAC.
                                                                                       (GLCNAC.
                                                                                                              (GLCNAC.
                                  (GLCNAC.
                 (GLCNAC.
                                              (GLCNAC.
                                                                      (GLCNAC.
                                                                                                   (GLCNAC.
                             (GLCNAC.
                                                                (GLCNAC.
     N-LINKED
                                                                                           N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                         CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 82.4
Best Local Similarity 77.5
Matches 158; Conservative
                                                                                                       2240
2529
2788
1035
1047
1050
1063
                                                                                                                                                            1085
1085
1087
1092
1097
1103
                                                                                                                                                                                                          1112
1120
1122
11133
11136
11161
11161
1171
1171
1174
11186
                                                                                       645
2041
2077
                                                                                                                                                                                                                                                                                                            1688
                                                                                                                                                      9101
                                                                                                                                                                                                                                                                                                                  Ř
                                                                                                                                                                                                                                                                                          1189
1198
1203
1680
3010
                                                                                                                                                1068
                                                                                                                                                                  1082
1086
1090
                                                                                                                                                                                   1093
                                                                                                                                                                                                                                                                              1175
                                                                                                                                                            1077
                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                         174
                                  CARBOHYD
CARBOHYD
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                 SEQUENCE
                  CARBOHYD
                        CARBOHYD
                             CARBOHYD
                                                                CARBOHYD
                                                                            CARBOHYD
                                                                                 CARBOHYD
             CARBOHYD
                                                                                                                              HELIX
STRAND
STRAND
STRAND
TURN
STRAND
                                                                                                                                                                              STRAND
                                                                                                                                                                                                               STRAND
STRAND
STRAND
TURN
STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                    STRAND
                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                              STRAND
                                                                                                                                                                                         STRAND
                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                          STRAND
                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                  HELIX
TURN
                                                                                                                                                                                                                                                             FURN
                                                                                                                                                                                                                                                                                                                                                                                                           g
g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE-91192160; PubMed-1849488;

MEDLINE-91192160; PubMed-1849488;

MEDLINE-91192160; PubMed-1849488;

MACON N. Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,

A Ohkoshi S., Shimotohno K.;

"Molecular structure of the Japanese hepatitis C viral genome.";

PEBS Lett. 280:325-328(1991).

"Molecular structure of the Japanese hepatitis C viral genome.";

PEBS Lett. 280:325-328(1991).

"I FUNCTION: THE SMALL PROFIEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

"S3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

"C ATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

"C ATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).

"I POPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN G AND MRNA.

"IPOPROTEIN C AND MRNA.

"SHILARITY: THE PROTESE BELOWGS TO PEPTIDASE FAMILY S29.
                                                                        01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-AG-1992 (Rel. 24). Last sequence update)
01-AG-1992 (Rel. 41, Last annotation update)
03-Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope qlycoprotein El (GP32) (GP35); Envelope glycoprotein E2 (GP88) (GP70) (NS1); Protease/helicase NS3 (P70) (Hepaclvirin)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepaclvirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P50); Nonstructural protein NS4B (P60) (P70) (RnA-directed RNA Pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Japanese) (HCV)
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                          PRT; 3010 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS1.
                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HSSP, P26663, 1JXP.
MEROPS, S29.001; -.
MEROPS; U39.001; -.
InterPro, IPR001410; E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                             Hepacivirus
                                        POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
RESULT 5
POLG_HCVJA
```

```
ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     M Polyprotein; Glycoprotein: Transferase; RNA-directed RNA polymerase;
M Core protein; Goat protein: Envelope protein; Helicase; APP-binding;
Transmembrane; Nonstructural representation of the protein of the prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 AAQIFLAICINGVCWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82.4%; Score 850.5; DB 1; Length 3010; 76.5%; Pred. No. 3.8e-72; Live 23; Mismatches 16; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3010 AA; 327017 MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL)
DECH BOX.
irPro; IPR002166; HCV_RGRP.

irPro; IPR001650; Helicase C.

erpro; IPR007095; RNA_pol_DS_PS.

ierpro; IPR007094; RNA_pol_DS_PS.

ierpro; IPR007094; RNA_pol_DS_ri.

iem; PF01542; HCV_core; 1.

iem; PF01542; HCV_core; 1.

iem; PF01542; HCV_core; 1.

iem; PF01543; HCV_mv; 1.

iem; PF01560; HCV_NS1; 1.

Pfam; PF02097; HCV_NS2; 1.

Pfam; PF02006; HCV_NS4; 1.

Pfam; PF01001; HCV_NS4b; 1.

Pfam; PF01001; HCV_NS4b; 1.

Pfam; PF01001; HCV_NS4b; 1.

Pfam; PF01001; HCV_NS4b; 1.

Pfam; PF01005; HCV_NSSp; 1.

Pfam; PF01005; HCV_NSSp; 1.

Pfam; PF01099; Viral_RGRP; 1.

Pfam; PF00998; Viral_RGRP; 1.

Probom; P1186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      best Local Similarity
Matches 156; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                     Solution of the state of the st
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-9229514: PubMed-1318627;

MEDLINE-92295714: PubMed-1318627;

MEDLINE-92295714: PubMed-1318627;

MEDLINE-92295714: PubMed-1318627;

MEDLINE-92295714: PubMed-1318627;

Molecular cloning of hepatitis C virus genome from a single Japanese carière: sequence variation within the same individual and among infected individuals.";

Virus Res. 23:39-53(1992).

Virus Res. 23:39-53(1992).

Virus Res. 23:39-53(1992).

C -- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMERANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

C -- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

C -- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                     01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein [GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (RC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin) NS4B (P27); Nonstructural protein NSAA (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); CP70) (RAM-directed RNA POLymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          {RNA}(N).
-!- SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus (isolate HC-JT) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                 PRT; 3010 A.A.
                 174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; D11168; BAA01943.1; -. PIR; A45573; A45573.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR007095;
IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=31642;
                                                                                                                               POLG_HCVJT
000269;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                     Hepaciv
                                                                                             RESULT
                                                                                                                                                       8
```

InterPro;

œ

```
54 AAQTFLATCINGVCWTVFHGAGIRTIASPKGPVIOMYINVDKDLVGWPAPQGSRSLIFFCT 113
                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81.8%; Score 844.5; DB 1; Length 3010; 77.5%; Pred. No. 1.4e-71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ?
.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94A1C77435D642BB
                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20; Mismatches
                                                                                                                                                                                                                                                                                                                             POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 CTRGVAKAVDFIPVESLEITMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645 645 N-LIN
2041 2041 N-LIN
2040 2240 2240 N-LIN
2529 2529 N-LIN
278 278 N-LIN
3010 AA: 326573 HW: C
Pfam; PF01543; HCV_capsid; 1. Pfam; PF01542; HCV_core; 1. Pfam; PF01550; HCV_core; 1. Pfam; PF01560; HCV_NS1; 1. Pfam; PF01960; HCV_NS2; 1. Pfam; PF01906; HCV_NS3; 1. Pfam; PF01001; HCV_NS4a; 1. Pfam; PF01001; HCV_NS5a; 1. Pfam; PF01501; HCV_NS5a; 1. Pfam; PF01501; HCV_NS5a; 1. Pfam; PF00571; helicase_C; 1. Pfam; PF00571; helicase_C; 1. Pfam; PF00598; Viral_RRRP; 1.
                                                                                                                             ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 77.5
Les 158; Conservative
                                                                                                                                                                                                                                                                                                                             369
11083
11107
11237
1319
1319
209
209
417
417
423
                                                                                                                                                                                                                        115
191
383
729
1006
1615
1862
2013
                                                                                                                                                                                       3D-structure
                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                 INIT_MET
                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
Matches
                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
```

4

InterPro; IPR002166; HCV_RdP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_bS_PS.
Pfam; PF01543; HCV_caps1d; 1.

HCV_NS5a

InterPro; IPR000745; InterPro; IPR001490; InterPro; IPR002868;

HCV_capsid.
HCV_core.
HCV_core.
HCV_NSI.
HCV_NSI.
HCV_NSI.
HCV_NSI.

InterPro; IPR002531;

InterPro; IPR001410; InterPro; IPR002522; InterPro; IPR002521; InterPro; IPR002519;

MEROPS; \$29.001; MEROPS; U39.001;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-1., Sugiyama Y., Kurai K., Lizuka H.,
Akchida A., Miyakawa Y., Mayumi M.;
Machida A., Miyakawa Y., Mayumi M.;
Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved of degree of the genomic RNA of the propertion of the sequence of the genomic RNA of the NA of 
                                                                                                                                                                                                                               01-AGC-1992 (Rel. 23, Created)
01-AGC-1992 (Rel. 23, Last sequence update)
01-AGC-1992 (Rel. 21, Last sequence update)
01-AGC-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NAS); Protean P7: Nonstructural protein NS2 (RP2);
(EC 3.4.22.-); Protease/hellcase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS54 (P4); Nonstructural protein NS58 (P65) (P70) (NNA-directed RNA polymerase) (EC 2.7.7.48)].
NS58 (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J6) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSIGTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                             PRT; 3033 AA.
1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303. HSSP; P27958; 1HEI.
                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus.
NCBL_TaxID=11113;
                                                                                     RESULT 7
POLG_HCVJ6
POLG_HCVJ6
?<66;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1034 TAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGTTISGVLWTVYHGAGNKT 1093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCICGSSDLYLVTRHADVIPVRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        78
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65.3%; Score 674; DB 1; Length 30:
68.2%; Pred. No. 2e-55;
ive 26; Mismatches 31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F957F5C1A273BE9E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GECNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DECH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AA; 329165 MW;
                                                                                                                                          Pfam: PF01006; HCV_NS4a; 1.
Pfam: PF01010; HCV_NS4b; 1.
Pfam: PF01506; HCV_NS5a; 1.
Pfam: PF00271; helicase_C: 1.
Pfam: PF00271; helicase_C: 1.
ProDom: P1186062; HCV_NS1; 1.
SMART: SMO0487; DEXDC: 1.
                                                         HCV_NS1; 1.
HCV_NS2; 1.
HCV_NS3; 1.
   HCV_core; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            122; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     116
192
384
734
734
1011
1620
1867
2018
347
                                                         PF01560;
PF01538;
PF01542;
                                                                                                                  PF02907;
                              PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                              pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
SOLA SELECTION OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
```

3033 AA

STANDARD;

POLG_HCVJ8 P26661;

82

POLG_HCVJ8

RESULT 8

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way one-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
               01-703 (Rel. 2), Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
38-FEB-2003 (Rel. 41, Last 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or The in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                   MEDIINE-92230232; PubMed-1314459; Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tsuda F., Mishiro S.; Eukuda S., Tsuda F., Mishiro S.; Full-length sequence of a hepetitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (RNA)(N).
-1- SUBMNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VITOLOGY 188:331-341(1992).

-1- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -1 - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002166; HCV_RGRP.
IPR007095; RNA_pol_DS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10988; BAA01761.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS5a; 1.
Viral_RdRP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01543; HCV_caps1d;
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS4a; 1.
HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_core; 1
HCV_env; 1.
HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A40250;
                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
```

```
DOMAIN
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HHOA_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LIPID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
       NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                     CELULAR AMINOPEPTIONSE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE./HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                        SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                     Core protein; Coat protein; Envelope protein; Helicase; AfP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Treponema pallidum.
Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
NCBI_TaxID=160;
                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65.1%; Score 672; DR 1; Length 3033; 68.7%; Pred. No. 3.1e-55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            330177 MW; 1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC
N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        485 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
  PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 196
209
233
299
305
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      649 64
1091 109
2038 203
2359 235
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           423
430
448
477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 9
Y136_TREPA
ID Y136_TREPA
AC 083172;
                                                                                                                                                                                                                                                                                                                                                      CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TP0136.
                                                                                                                                                                CHAIN
                                                                                                                                                                                                               CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                CHAIN
  NAME OF A STATE OF THE STATE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 윱
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 AGSKLYATNGRLWEKELNGTGSWQRVSSSSVPTDSDK----KVMSIATDGNTFVLACVP 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   64 -NGVCWTVYHGAG---TRIIASPKGPVIQMYINVDKDLVG-----WPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 C----GSSDLYLVTRHADVIP-----151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CGGGGGGSSSSSSSCIHIMLVPGGTGNNGNCGCGGGGGSSSSSSCIHIKVENTDEQFL 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         109 GTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT----LVGGTSKPFWLVPGGTGNNGNCG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SGDTAYA-----QQTRGEEGCQETSQ----TGRDKNOVEGEVQIVSTAAQTFLATCI- 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HHOA_ARATH STANDARD; PRT; 321 AA.
09SEL7: 049507;
16-027-2001 (Rel. 40, Created)
16-027-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HhoA, chloroplast precursor (EC 3.4.21.-).
HHOA OR AT4G18370 OR F28312.30.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                71; Gaps
                                                                                            Fraser C.M., Norits S.J., Weinstock G.M., White O., Sutton G.G., Doddson R., Gwinn M., Bickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujil C., Garland S., Wenter J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                 Science 281:375-388(1998).
-!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
                                                                                                                                                                                                                                                                                                                                                  "Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYPOTHETICAL LIPOPROTEIN TP0136.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (Potential)
-!- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 85.5; DB 1; Length 485; Pred. No. 1.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-ACYL DIGLYCERIDE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----GSSGGPLLCPAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      225 DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           $; Pred. No. 1.3;
17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLY/SER-RICH.
GLY/SER-RICH.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLY/SER-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; AE001199; AAC65137.1; ALT_INIT.
                             STRAIN=Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       48984 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            50; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Complete proteome
```

```
Nature 402:769-777(1999).
     SEQUENCE FROM N.A
```

```
MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 321 AA; 34691 MW;
                                                                                                                                                                                                                                                                                                                            Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 VDF-IPVESLETTM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |:| ||:::: |:
294 VNFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
les 56; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAMP_HUMAN
Q13685;
                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
DOMAIN
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                               TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                           A STRAIN-C. Columbia.

AND COLUMBLE TOOL UNDIA.

AND COLUMBLE TOOL UNDIA.

AND COLUMBLE TOOL UNDIA.

AND COLUMBLE TOOL UNDIA.

AND COLUMBLE TOOL WAS BRANCH TO TO THE TOOL TO THE TOOL 
                                                                                     SEQUENCE FROM N.A.
Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease, a homolog to the bacterial periplasmic protease HhoA.";
submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320.
Schubert M., Peterson U., Funk C., Haas B., Schroeder W.P.,
Kleselbach T.;
eurosids II; Brassicales: Brassicaceae; Arabidopsis.
NCBL_TaxID=3702;
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its

13; use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch). 140 ----SLLSPRPISYLK-----GSSGGPLLCPAGHAVGIFRAAVCTR---GVAKA 181 235 SGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSYGHTIGV-NTATFTRKGSGMSSG 293 95 KISPSVVVIEAIELPKISSGDIL----- 129 2 KKKGSVVIVGRINL----SGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQT 57 58 FLATCINGVCW-----TVYH----TVYH-----GAGIRTIASPKGPVIQM 89 182 -- DPDNDLAVLKIETEGREINPVVLGTSNDLRVGQSCFAI-----GNPYGYENTLTIGVV 69; Indels 102; Gaps 90 YINVDKDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSRG------Chloroplast; Thylakoid; Iransit peptide. CHLOROPLAST (POTENTIAL). Homo sapiens (Human).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606; CHARGE RELAY SYSTEM (POTENTIAL). CHARGE RELAY SYSTEM (POTENTIAL). CHARGE RELAY SYSTEM (POTENTIAL). 8.2%; Score 85; DB 1; Length 321; 22.0%; Pred. No. 0.93; -> G (IN REF. 1). 68DB81E0BD27A7A7 CRC64; 15-JUL-1998 (Rel. 36, Created) 15-JUL-1998 (Rel. 36, Last sequence update) 28-FEB-2003 (Rel. 41, Last annotation update) Angio-associated migratory cell protein. Conservative 27; Mismatches 452 AA PROTEASE HHOA. POLY-GLU. SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. THYLAKOID InterPro: IPR001940; Protease2C.
InterPro: IPR001254; Ser_protease_Try.
Pfam: PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASES2C. EMBL; AF114386; AAF24060.1; EMBL; AL021710; CAA16717.1; ALT_SEQ. EMBL; AL161548; CAB78839.1; ALT_SEQ. MEROPS; S01.279; PRT;

```
Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            104
152
324
171
201
280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydrolase;
TRANSIT
TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
       $$\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{2}\frac{1}{
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profil institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               φ,
.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    68 WIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWPAPQGSRSL----IPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA------ 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241 -----VLPDGKRAVVGYEDGTIRIMDLKQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             022609: 09LK85;
16-077-2001 (Rel. 40, Created)
16-077-2001 (Rel. 40, Last sequence update)
16-077-2001 (Rel. 41, Last annotation update)
Protease Do-11ke 1, chloroplast precursor (EC 3.4.21.-).
DGGP1 OR DGGP OR A73627925 OR K16N12.18.
Arabidopsis thallana (Mouse-ear cress).
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Hagnoliophyta: eudicotyledons: core eudicots; Rosidae: MCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            197 WMEWH-----PRAPVLLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR----
       IN ENDOTHELIAL CELLS, CYTOTROPHOBLASTS, AND POORLY DIFFERENTIATED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982: PubMed-9507020;
Itzhaki H., Naveh L., Lindahi H., Cook M., Adam Z.;
Ildentification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              163 ------GHAVGIFR----AAVCTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7.9%; Score 82; DB 1; Length 452; 25.3%; Pred. No. 2.7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13; Mismatches 47; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HEPARIN-BINDING (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DA1413D25EB236C0 CRC64;
                                  COLON ADENCARCINOMA CELLS FOUND IN LYMPHATICS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MIM; 603488; -.
GO: GO: 0008201; F: heparin binding activity; TAS.
Interpro: IFR001680; WD40.
SMART; SM00320; WD40; 8.
PROSITE; PS00063; WD.REPEATS_1; 1.
PROSITE; PS50082; WD.REPEATS_2; 6.
PROSITE; PS50294; WD.REPEATS_2; 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              437 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY -GLU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M95627; AAA68889.1; -. PIR; I39383; I39383.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18
77
138
180
220
220
261
306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genew; HGNC:18; AAMP.
MIM; 603488; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      452 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Repeat; WD repeat.
DOMAIN 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEG1_ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEG1_ARATH
   STEETETEE SARAHARA STEETETEETE STEETE STEETE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPRESENTATION OF THE PROPERTY OF THE PROPERTY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the SWISS Institute of Bioinformatics and the BMB. Outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ----PA 102
                                                                                                                                                                                                                                                                                     STRAIN-CV. Columbia:
Kieselbach T., Bystedt M., Schroeder W.P.;
Kleselbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
--- HAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
CAN DEGRADE BETA-CASEIN.
--- ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL, FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Transit peptide; Chloroplast; Thylakoid. CHLOROPLAST (POTENTIAL). THYLAKOID.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF.
SEQUENCE FROM N.A.
STRAIN-cv. Columbia;
MEDLINE-20363099; PubMed-10907853;
KADEVO T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
Kaneko T., Ratoh T., Sato S., Sato S., Seguenca A., Samizu E., Tabata S.;
Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 Pl. TAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -! - SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7.8%; Score 80.5; DB 1; Length 437; 25.6%; Pred. No. 3.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       55; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ---TNVDKDLVGW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       =
-
-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1497B1AB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2).

V -> I (IN REF. 2).

P -> S (IN REF. 2).

G -> D (IN REF. 2).

LL -> HF (IN REF. 2).

L -> V (IN REF. 2).

I -> V (IN REF. 2).

I -> V (IN REF. 2).

Q -> E (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEASE DO-LIKE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.-i- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SERINE PROTEASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     70 VYHGAGTRIIASPKGPVIQMY-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; 1901.07)
InterPro; 1PR001408; PDZ.
InterPro; 1PR001940; Protease2C.
InterPro; 1PR00154; Ser_protease_Try.
Pfam; PF00555; PDZ; 1.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00884; PROTEASES2C.
SMART; SW002284; PDZ; 1.
PROSITE; PS50106; PDZ; 1.
Hydrolase; Serine protease; Transit pept
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AP000371; BAB02539.1; --
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF028842; AAC39436.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -
::
<del>:</del>:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- INDUCTION: By heat shock.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 25.6 es 44; Conservative
                                                                                                                                                                                and BAC clones.";
DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       103
321
421
171
201
280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       O-PHENANTHROLINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ..
--
-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         416 4
437 AA;
                                                                                                                                                                                                                                                                   SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S01.279; -.
```

181 FVVARILNT 189

150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209

```
CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GK; Q02224;
MIM; 117143;
                                         CENE_HUMAN
                                                                                                                                                                                                                                                                   mitosis
          RESULT 14
CENE_HUMAN
                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  75 QIRVFGQND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALK 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   132 RRRGDSRGSLLSPR--PIS-----YLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVD 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           86 VIQMYINVDKDLVGWPAPQGSRSLTP-----CTCGSSDL-----YLVTRHADVIPV 131
                          210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
103 POGSRSLTPCTCGSSDLYLV------TRHADVIPVRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               29 QEEREVHFLISKAAQIVMAT------ETDVALPAKPQAMQAFLTEYCGAAAG 74
                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
STRAIN-ATCC 15692 / PAO1;
MEDLINE-2043737; Pubmed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Was A., Larbig K., Lim R.M.,
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen IT.,
Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   43 QVEGEVQ-IVSTAAQTFLATCINGVCWTVYHGAGTRTIASPKGP----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                      148 SYL------KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF-IPVESL 190
                                                                     266 QDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGASSGVGFSIPVDIV 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          67;
                                                                                                                                                                                                                                                                  Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIR; H83144; H83144.
InterPro; IPR003382; Flavoprotein.
Pfam: PF02441; Flavoprotein; 1.
Hypothetical protein; Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MW; 01PD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nature 406:959-964(2000).
-1- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE / PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7.6%; Score 78.5; DB 1; Length 209; 26.5%; Pred. No. 2.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                               16-CCT-2001 (Rel. 40, Created)
16-CCT-2001 (Rel. 40, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                      209 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                  Pseudomonadaceae; Pseudomonas
                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               opportunistic pathogen.
                                                                                                                                                                                                                                                       Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           184 FIPVESLET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
les 50; Conserv
                                                                                                                                                                                                                                                                                               NCBI_TaxID-287;
                                                                                                                                                       PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                        ö
                         셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
MEDLINE-9304922; PubMed-1406971;
MEDLINE-1304922; PubMed-1406971;
FOR T.J. Li G., Schaar B.T., Szilak I., Cleveland D.W.;
"CENP-E is a putative kinetochore motor that accumulates just before
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -1- SUBUNIT: INTERACTS WITH CENP-F AND BUBRI KINASE.
-1- SUBCULT. INTERACTS WITH CENP-F AND BUBRI KINASCHORES DURING
CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND QUANTITATIVELY DISCARABED AT THE END OF THE CELL DIVISION.
-1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.; "Mitotic HeLa cells contain a CENP-E-associated minus end-directed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHÁRACTERIZATION.

CHÁRACTERIZATION.

CHADLIKE-94843747. PubMed-9763420;

Chan G.K.T., Schaar B.T., Yen T.J.;

Characterization of the kinetochore binding domain of CENP-E reve

"Characterization of the kinetochore proteins CENP-F and hBUBRI.";

J. Cell Biol. 143:49-63(1998).

-I- FUNCTION: MINUS-END DIRECTED MICROTUBGLE MOTOR. PROBABLE

RINETOCHORE MOTOR. ACCUMULARES JUST BEPORE MITOSIS AT THE G2 P

OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME NOVEMENT

AND/OR SPINDLE ELONGATION.
                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     60; G0:0005699; C:Kinetochore; TAS.
G0: G0:0005634; C:Holeleus; TAS.
G0: G0:00008350; F:Kinetochore motor activity; TAS.
G0: G0:000067; F:Kinetochore motor activity; TAS.
G0: G0:0000067; P:DNA replication and chromosome cycle; TAS.
G0: G0:0007079; P:mitotic chromosome movement; TAS.
G0: G0:0007089; P:mitotic metaphase plate congression; TAS.
InterPro; IPR001752; Kinesin_motor.
                                                                                    01-JUL-1993 (Rel. 26, Created)
01-JUL-1993 (Rel. 26, Last sequence update)
01-MAY-2000 (Rel. 39, Last annotation update)
Centromeric protein E (CENP-E protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRINTS; PR00380; KINESINHEAVY.
SWART; SM00129; KISC; 1.
PROSITE; PS00411; KINESIN_MOTOR_DOMAIN; 1.
PROSITE; PS50067; KINESIN_MOTOR_DOMAIN2; 1.
2663 AA
    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDLINE-95196755; PubMed-7889940;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; 215005; CAA78727.1; -. PIR; S28261; S28261.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam; PF00225; kinesin; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    microtubule motor.";
EMEO J. 14:918-926(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMD., 2.
PIR; S28261; S28261.
HSSP; P17119; 3KAR.
Genew; HGNC:1856; CENPE.
    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nature 359:536-539(1992)
                                                                                                                                                                                                                                                                                                        Homo sapiens (Human)
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                             2523 QONEQLIKOKNELLSNNQHLSNEVKTWKERTLKREAHKOVTCEN-------2566
                                                                                                                                                                                                                                                                                                               2567 SPKSPKVTGTASKKK------QITPSQCKERNL-----QDPVPKESPKSCFFD 2608
                                                                                                                                                                                                                                                                                81 SPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRG----D 136
                                                                                                                                                                                                            32 QETSQTGRDKNQVEGEVQIVSTAAQTF-----LATCINGVCWTVYHGAGTRTIA 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
STRAIN-CCDG 37608 / M982 / Serogroup B / Serotype 9;
MEDLINE-9334825; PubMed-8344530;
MEDLINE-934825; PubMed-8344530;
Jacoba E., Schryvers A.B.;
                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                      2609 SRSKSLPSPHPVRYFDNSSLG--LCPEVQNAG------AESVDSQP 2646
                                                                                                                                                                                                                                                                                                                                                  SRG-SLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIP 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -1 - FUNCTION: ACTS AS A TRANSFERRIN RECEPTOR AND IS REQUIRED FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TRANSFERRIN UTILIZATION.
-1- SUBCELLULAR LOCATION: Outer membrane.
-1- INDUCTION: By iron starvation.
-1- SIMILARITY: LOCAL TO OTHER TONB-DEPENDENT RECEPTOR PROTEINS.
Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;
                                                                                                                                                                           63;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Cloning and characterization of Neisseria meningitidis genes encoding the transferrin-binding proteins Tbpl and Tbp2."; Gene 130:73-80(1993).
                                                                                                                                         DB 1; Length 2663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Neisseria meningitidis (serogroup B).
Badteria; Proteobacteria: Betaproteobacteria; Neisseriales;
Neisseriaceae; Neisseria.
NCBL_TaxID-491;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSFERRIN-BINDING PROTEIN 1.
                                                                                                                                                                           52; Indels
                                                     COILED COIL (POTENTIAL).
GLOBULAR (POTENTIAL).
ATP (BY SIMILARITY).
MW: CEFC13880C6C8CB8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                38 45 TONB BOX.
894 911 TONB C-TERMINAL BOX.
911 AA; 101631 MW; 99283ABAE0B773E6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Transferrin-binding protein 1 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         911 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PIR; JN0821; JN0821.
InterPro; IPR000531; TonB_boxC.
Pfam; PF00593; TonB_dep_Rec; 1.
PROSITE; PS00430; TONB_DEPENBENT_REC_1; 1.
PROSITE; PS01156; TONB_DEPENBENT_REC_2; 1.
Outer membrane: Receptor: Signal; TonB_box.
                                                                                                                                                                           15; Mismatches
                                                                                                                                                            Pred. No. 56;
                                      KINESIN-MOTOR
                                                                                                                                         7.5%; Score 77.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                          2663 AA; 312087 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; Z15130; CAA78833.1; -.
                                                                                                                                                            24.0%;
                                                                                                                                                                            41; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                      335
2471
2663
                    Cell cycle: Centromere.
DOMAIN 1 335
                                                                                                                                                            Similarity
                                                                     2472
86
                                                     336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TB11_NEIMB
Q09056;
                                                                     DOMAIN
NP_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                           Query Match
Best Local
                                                         DOMAIN
                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SETTT: S
                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                    õ
```

٠.

```
13;
                                                                                                                                                                                                                                                                                                                                                        155 ----KGSNSVEQ---GSGALAGSVAFQTKTADDVIGEGRQWGIQSKTAYSGKNRGLTQS- 206
                                                                                                                                          :: | :|:: :|| :||:
50 RRDNEVTGLGKLVKTADTLSKEQVLDIRDLTRYDPGIAVVEQGRGASSGYSIRGMDKNRV 109
                                                                                                                                                                                                                                                            110 SLIVDGLAQIQSTIAQAAL ------GGTRI-AGSSGAINEIEYENVKAVEIS 154
                                                                                                         2 KKKGSVVIVGRINLSGDTAYAQQ-----TRGEEGCQETSQ------TGRDKNQ- 43
                                                                                                                                                                                                           44 ---VEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQM-YTNVDKDLVG
                                                                                                                                                                                                                                                                                                             100 WPAPQGSRSLTPCTCGSSDL----YLVTRHADVIPVRR------GDSRGSLLSP
                                                        Gaps
                                                        :99
                                                                                                                                                                                                                                                                                                                                                                                                                 145 RPISYLKGSSGG--PLLCPACHAVGIFRA-AVCTRGVAKAVDFIPVE 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---IALAGRIGGAEALLIHTGRRAGEIRAHEDAGRGVQSFNRLVPVE 250
7.4%; Score 76; DB 1; Length 911; 26.4%; Pred. No. 22; tive 20; Mismatches 81; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        completed: August 30, 2003, 19:13:45
he : 10.7567 secs
     Query Match 7.4%
Best Local Similarity 26.4%
Matches 60; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search con
Job time
                                                                                                                                                       q
                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                  ò
```

hepatitis control of the patitis control of t

hepatitis hepatitis hepatitis

hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis
hepatitis

090183 090186 090000 090000 090188 090188 090188 090188 090188 090189 090189 090189 090189 090189 090189 090189

091RS3 0950HG6 0950HG6 099FW19 099FW19 0991RS4 0991RS4 0991RS4 0991RR4 0991RR4 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7 0991RR7

```
SEQUENCE FROM N.A.
MEDLINE-2013484; PubMed-10864644;
MEDLINE-2013484; PubMed-10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-
dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lai V.C.H., Hong Z.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL: AF568278; AAF82566.1; -.
HSSP: P26663; LJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
Mucosal disease virus.
Viruses; ssRNA positive-strand viruses,
Pestivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; S31.001; -.

Interpro: IPR000180; CDVir_endptseP80.

Interpro: IPR001109; BCAD.

Interpro: IPR001109; HCV_NS3.

Interpro: IPR0010569; HGV_RGRP.

Interpro: IPR0010559; HGV_BAD.

Interpro: IPR0010559; MYD_DNA_Dinding.

Interpro: IPR001058; RNA_DOl_DS_PS.

Interpro: IPR007094; RNA_DOl_DS_PS.

Interpro: IPR007094; RNA_DOl_PSVIr.

Pfam; PF02907; HCV_NS3; 1.

Pfam; PF02907; HCV_NS3; 1.

Pfam; PF00998; Viral_RGRP; 1.
  PRELIMINARY;
 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WCBI_TaxID-11099;
 Q9IFH8
                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 1
Q91FH8
 DDAR REPARED DE LA DESTRETA DE LA DELLA DE LA DELLA DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091fh8 mucosal dis 081756 hepatitis c 09365 hepatitis c 0036579 hepatitis c 0931rT hepatitis c 091rr3 hepatitis c 091rr3 hepatitis c 091rr1 hepatitis c 091rr1 hepatitis c 091rr1 hepatitis c 091rr2 hepatitis c 091rr3 hepatitis c 091rr5 hepatitis c 091rr2 hepatitis c
                                                                                                        (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                         August 30, 2003, 19:00:22; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                        1032
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                       830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                         summaries
                                                                  OM protein – protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091FH8
081756
0036579
0036579
091RR8
091RR5
091RR3
091RR3
091RR0
091RR1
091RR3
                                                                                                                                                                                                BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                            sp_archea:*
sp_bacteria:*
sp_fung:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_plant:*
sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                          Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2222222222222222222222222
                                                                                                                                            US-09-965-594-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                  SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4040
2436
3011
3011
3011
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
181
181
181
181
3011
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 910.5
901.5
901.5
900.5
896.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   892.5
891
891
890
                                                                                                                                               Title:
Perfect score:
                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        893
893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
                                                                                                                                                                         Sednence:
                                                                                                                                                                                                                                        Searched:
                                                                                                                                                                                                                                                                                                                                                                                  Database
                                                                                          on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
```

no DNA stage; Flaviviridae;

PRT; 4040 AA

ALIGNMENTS

~

```
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sednence.
                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            09IFE5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
Q91FE5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEATH TO DEATH THE TENT TO THE
                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 CINGVCWTVYHGAGTRTIASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNOVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLAI 61
PRINTS; PR00729; C.C.
SWART; SM00407; DEXDC; I.
SWART; SM00407; DEXDC; I.
SWART; SM00409; HELICC; I.
PROSITE; PS00037; HYB 1; I.
PROSITE; PS00501; RDRP_POSITIVE; I.
PROSITE; PS00531; RNAS_T2_2; I.
ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase.
**Content of the content of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match

88.2%; Score 910.5; DB 12; Length 4040;
Best Local Similarity 91.3%; Pred. No. 4.5e-83;
Matches 178; Conservative 5; Mismatches 9; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Choo Q.-L., Richman K., Han J.;
Choo Q.-L., Richman K., Han J.;
The nucleotide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL; M32084; AAA45677.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          081756 PRELIMINARY; PRT; 2436 AA.
081756;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
denome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD.
InterPro; IPR001251; HCV_NS1.
InterPro; IPR0012518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001409; HCV_NS4a.
InterPro; IPR001466; HCV_NS4b.
InterPro; IPR001666; HCV_NS5a.
InterPro; IPR001666; HCV_NSSa.
InterPro; IPR001669; RNA_PO1_DS_PS.
InterPro; IPR007094; RNA_PO1_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfan; PF01560; HCV_NS1; I. Pfan; PF01589; HCV_NS2; I. Pfan; PF0200; HCV_NS3; I. Pfan; PF01006; HCV_NS4s; I. Pfan; PF01001; HCV_NS4b; I. Pfan; PF01506; HCV_NS5s; I. Pfan; PF01506; HCV_NS5s; I. Pfan; PF00271; helicase_C: I. Pfan; PF00998; Viral_RGRP; I. Probom; PP186062; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         190 VDFIPVENLETTTRS 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRANCE OF THE PROPERTY OF THE 
                          DR DR DR COR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
```

614 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173 23 3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST SEQUENCE FROM N.A. MEDINE-11369872; MEDINE-2122: Pubmed-11369872; Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.; Infectious cDNA clone of the hepatitis C virus genotype 1 prototype PROSITE; PSS0521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Hellcase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transmembrane. Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus. NCBL_TaxID=11103; J. Geo. Virol. 82:1291-1297(2001).
-!- SUBUINT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; AF271632; AAF81759.1; -HSSP; P27958; 1AIV. 6 87.4%; Score 901.5; DB 12; Length 2436; 85.8%; Pred. No. 1.9e-82; tive 9; Mismatches 11; Indels 9; 2436 2436 2436 AA; 264734 MW; D7B9B72900BE3125 CRC64; 01-OCT-2000 (TrEMBLrel. 15, Created) 01-OCT-2000 (TrEMBLrel. 15, Last sequence update) 01-WAR-2003 (TrEMBLrel. 23, Last annotation update) PRT; 3011 AA nterPro; IPR000345; CytC_heme_bind.
InterPro; IPR001410; DEAD.
InterPro; IPR001410; DEAD.
InterPro; IPR001521; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_MS1.
InterPro; IPR004519; HCV_MS2.
InterPro; IPR004409; HCV_MS3.
InterPro; IPR001490; HCV_MS4.
InterPro; IPR001490; HCV_MS4.
InterPro; IPR001490; HCV_MS4.
InterPro; IPR001490; HCV_MS4.
InterPro; IPR001490; HCV_MS5.
InterPro; IPR001509; HCV_MS5.
InterPro; IPR001509; HSIQASe_C.
InterPro; IPR001509; RNA_DOl_DS_PS.
InterPro; IPR007094; RNA_DOl_DS_PS.
IPR0M; PF01542; HCV_CORP; I. 174 CTRGVAKAVDFIPVESLETTMRSP 197 Best_Local Similarity 85.8 Matches 175; Conservative PRELIMINARY; Genome polyprotein. Hepatitis C virus.

```
SEQUENCE FROM N.A.
                                                                                   SEQUENCE FROM N.A.
                                                                                                         STRAIN-HC-J1;
                                                                                                                                                                                                                STRAIN-HC-J
                                                                                                                             Okamoto H.;
  SOW WE WE WANTED THE WAY WANTED TO THE WAY WANTED THE WANTED THE WAY WANTED THE WANTED THE WAY W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               용
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 AAQTFLATCINGVCWTVYHGAGTRIIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                          53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                    PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS05097; RDRP_POSITIVE; 1.
ATP-Dinding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydroclase; Nonstructural protein; Pobyprotein;
Hydroclase; Nonstructural protein; Pobyprotein; SRNA-directed RNA, polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 MM; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NNV-1908 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
Heparitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                           DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "The 5'-terminal sequence of the hepatitis C virus genome."; Jpn. J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-HC-JJ;
MEDLINE-91013116; Pubmed-2170712;
Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                     11; Indels
                                                                                                                                                                                                                                                                                                                                                         87.4%; Score 901.5; DB 1:
85.8%; Pred. No. 2.5e-82;
iive 9; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-HC-J1;
MEDLINE-93117120; PubMed-1335573;
Okamoto H., Kanal N., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-HC-J1;
MEDLINE-92044440; PubMed-1658196;
                     Pfam; PF01005; HCVAS4a; 1.
Pfam; PF01001; HCVAS4a; 1.
Pfam; PF01001; HCVAS5a; 1.
Pfam; PF00271; hcVAS5a; 1.
Pfam; PF00398; Viral_RGRP; 1.
Probom; PD186062; HCVAS1; 1.
                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 85.8
Matches 175; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               003463;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           003463
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         003463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ô
```

```
Ϊ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1005 RKGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 AAQIFLATCINGVCWTVFHGAGTRIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLIPCI 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydroclase; Nonstructural protein; Polyprotein; Rolyprotein; Rolyprotein; Rolyprotein; Rolyprotein; Rolyprotein; Stansmembrane. SEQUENCE 3011 AA: 327112 MW; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                 "Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-JI) with high homology to USA isolates."; Nucleic Acids Res. 20:6410-6410(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                      Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12; Indels
                                                                                                                                                                                                   Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 900.5; DB 1 Pred. No. 3.2e-82;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 87.3%; Score 900.5; ,
Best Local Similarity 85.8%; Pred: No. 3.2e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; INROUGES; Helicase_C.
Interpro; IPROUGES; Helicase_C.
Interpro; IPROUGOS; RNA_pol_DS_PS.
Interpro; IPROUGOSS; RNA_pol_DS_PS.
Interpro; IPROUGOSS; RNA_pol_DS_PS.
Interpro; IPROUGOSS; INPO_COTE; I.
Pfam; PFOLIS43; HCV_core; I.
Pfam; PFOLIS50; HCV_NS; I.
Pfam; PFOLIS50; HCV_NS; I.
Pfam; PFOLIS50; HCV_NS4; I.
Pfam; PFOLISOS; HCV_NS4; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART; SM0487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HSSP; P4/298; Anti.
Interpro; IPR001410; DEAD.
Interpro; IPR001521; HCV_capsid.
Interpro; IPR002521; HCV_capsid.
Interpro; IPR002519; HCV_core.
Interpro; IPR002519; HCV_MS.
Interpro; IPR002518; HCV_MS.
Interpro; IPR001409; HCV_MS.
Interpro; IPR001409; HCV_MS.
Interpro; IPR001409; HCV_MS.
Interpro; IPR00140; HCV_MS.
Interpro; IPR00140; HCV_MS.
Interpro; IPR00166; HCV_MS.
Interpro; IPR00166; HCV_MS.
                                                                                                                                                                                                                                                                                                               MEDLINE-94174722; PubMed-7510436;
```

us-09-965-594-14.rspt

```
54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-Pt.1Y;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       091RT5
                                                                                                                                                                                                                                                                                                                                                                                    091RRB;
                                                                                                                                                                                                                                                                                                                RESULT 6
                                                                                                                                                                                                                                                                                                                                        091RR8
                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                               q
                                                                                                              ò
                                                                                                                                                   q
                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1005 RRGQEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-97373636; PubMed-9228008;
Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone S.
                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86.9%; Score 896.5; DB 12; Length 3011; 84.8%; Pred. No. 8.2e-82; ive 10; Mismatches 12; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydrolase, Monstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327182 MW; E2E0EE809C63C1B9 CRC64;
                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001410; DEAD.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002531; HCV_core.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR002519; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001650; HCV_NS4.
InterPro; IPR001650; HCV_NS4.
InterPro; IPR001690; HCV_NS4.
InterPro; IPR001690; HCV_NS4.
InterPro; IPR001690; HCV_NS4.
InterPro; IPR001690; HCV_NSA.
InterPro; IPR001690; HCV_NSA.
InterPro; IPR001690; RNA_pol_DS_PS.
InterPro; IPR001690; RNA_pol_PSvir.
Pfam; PF01543; HCV_CODE; 1.
                     1185 CTRGVAKAVDFIPVESLETTMRSP 1208
CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                       Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PP01001; HCV_NS45; 1.
Pfam; PP01506; HCV_NS5a: 1.
Pfam; PF001505; HCV_NS5a: 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1: 1.
                                                                                                                                                                                                  01-JAN-1998 (TrEMBLrel. 05, 01-JAN-1998 (TrEMBLrel. 05, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 173; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam; PF01538; HCV_NS2; 1
Pfam; PF02907; HCV_NS3; 1
Pfam; PF01006; HCV_NS4a;
                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                      Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF01560; PF01538; 1
                                                                                                                                                                                                                                                                                                                                        Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rice C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                          036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                              RESULT :
                                                                                                                                                          SO WEW WITH THE PROPERTY OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                            셤
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19 TAYAQQIRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQIFLAICINGVCWTVYHGAGIRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCMTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-201) to the EMBL/GenBank/DDBJ databases.
EMBL, AFS9235; AAK54560.1; -
InterPro: IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NS3 protease (Fragment).
Hepatitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepativirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protease.

NON_TER 181 181

SEQUENCE 181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 23, Last annotation update)
NS3 professe (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        86.8%; Score 896; DB 12; 96.6%; Pred. No. 2.5e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. 2.5e
1; Mismatches
                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. STRAIN-Pt.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAGE 11 PAGE 1
```

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAIN-Pt. 4B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-Pt.K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Protease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
Q91RS1
                                                     291RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SO DE RELEGIO DE LA CONTRE LA CONTRE
                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCTGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 TAYAQQTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFRN of the NS3 Protease Gene from
"Clinical Strains of the Hopatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/ODBJ databases.
EMBL, AF862538, ARK54631:
InterPro; IPR004109; HCV_NS3.
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369218; AAK54543.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 896; DB 12; Length 181;
Pred, No. 2.5e-83;
1; Mismatches 5; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 894; DB 12; Length 181;
Pred. No. 4e-83;
2; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 181
181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                               181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last seq
01-MAR-2003 (TrEMBLrel. 23, Last ann
MS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                      86.8%;
96.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match 86.6%;
Best Local Similarity 96.1%;
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 96.68
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Pt.3U;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Protease.
                                                                                                                                                                                                                                                                              Protease.
                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local (
                                                                                                                                                                                                                                                                                                            NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT RE
```

g

õ

g

à

g

g

ογ

ð 8 ò

```
78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4 TAYAQOTRGLIGCIITSLIGRDKNQVBGEVQIVSTAAQTFLATCINGVCWTVYBGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                         Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to Irw of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AFS62401, ARK5455.1;
Interpro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2011) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369232; AAK4557.1; -.
Interpro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                      Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
Last sequence update)
Last annotation update)
                    091RR3;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86.5%; Score 893; DB 12; 96.1%; Pred. No. 5.1e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 893; DB 12;
Pred. No. 5.1e-83;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 5.1e
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            86.5%;
96.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel, 19, 01-DEC-2001 (TrEMBLrel, 19, 01-MAR-2003 (TrEMBLrel, 23, NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 181 181 181 181 181 181 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
```

a

g õ

```
Desai S.M., Devare S., Yamaguchi J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_DSvir.
Pfam; PF01543; HCV_capaid; 1.
Pfam; PF01539; HCV_cenv; 1.
Pfam; PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro, IPR000345; CytC_heme_bind
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR000745; HCV_NS4a.
Interpro; IPR001499; HCV_NS4b.
Interpro; IPR0021669; HCV_NS5a.
Interpro; IPR002166; HCV_RGRP.
Interpro; IPR001650; Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 96.1:
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C Virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         P27958; 1HEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-colonel
                                                                                                                                                                                                                                                                                                                              NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                             Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9ELS8
Q9ELS8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DERENA OCCONDENSA DE LA CONTRE LA CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            79 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 IASPKGPVIQMYTNVDKDLVGWPAPQCTRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                       | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 1974 | 
                                                                                                                                                                                                                  IASPKGPVIQMYINVDKDLVGWPAPQGSRSLIPCICGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF1PVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-PL.SI;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369245; AAK54570.1;
InterPro; IPR004409; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  091R08
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 proteases (Fragment).
Hepatitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 181;
   5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 181
181 AA; 19144 MW; COC91F1E2EEB0B32 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-DEC-2001 (TIEMBLTEL. 19, Created)
01-DEC-2001 (TIEMBLTEL. 19, Last sequence update)
01-DEC-2003 (TIEMBLTEL. 23, Last annotation update)
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86.5%; Score 893; DB 12;
96.1%; Pred. No. 5.1e-83;
tive 2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 AA
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          171;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                              13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q91RT1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           091RT1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                091RQ8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
Q91RT1
```

SO OR REPARENTED TO THE SO OR OR SO OR SO

g ð g ò

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79 IASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 IASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCTCGSSDLYLVTRHADVIPVRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19 IAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUBMITTED (JUL-2000) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AFS65221, AARS6447.1;
-Interpro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

, AF290978; AAG02099.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                              Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              5; Indels
                                                                                                                                                                                                                                                                                      181 181
181 AA; 19114 MW; ABB90B5B3ABA4E26 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                           86.5%; Score 893; DB 12;
96.1%; Pred. No. 5.1e-83;
iive 2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3011 AA
```

```
NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                            091RS9;
                                                                                                                                                            091RS9
                                                                                                                                RESULT 15
Q91RS9
                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                               a
                                                             ò
                                                                                       8
                                                                                                                                                                                                                                                                                                                        1065 ATQTELATCINGVCWTVYHGAGTRIASPKGPVIQMYINVDQDLVGWPAPQGSRSLIPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                    1125 CGSSDLYLVTRHADVIPVRRQGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                               54 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 TAYAQQTRGLLGCIITSLTGRDKNQVEGELQIVSTAAQTFLATCINGVCWTVFUGAGTRT 63
                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-PL.3T;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369237; AAK54562.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3.
                                                                                                              PROSITE, PRODUGGT, DEXDE; 1. PROSITE, PRODUGGT, DEXDE; 1. PROSITE, PRODUGGT, CTCCHRONE_C; 1. PROSITE, PSO503; RDRP_POSITIVE; 1. ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; Polyprotein; RAA-directed RNA polymerase; Transferase; Transmembrane. SEQUENCE 3011 AA: 327107 MM; AGBECFSA3B3EE13F CRC64;
                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                               DB 12; Length 3011;
                                                                                                                                                                                                                                                                            ٠.
م
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 86.3%; Score 891; DB 12; Length 1
Best Local Similarity 95.5%; Pred. No. 8.2e-83;
Matches 170; Conservative 3; Mismatches 5; Indels
                                                                                                                                                                                                                                               86.5%; Score 892.5; DB 12; Length
84.3%; Pred. No. 2.1e-81;
tive 11; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 181
181 AA; 19101 MW; 614ADA8BOF33CCAF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
Pfam; PF01538; HCV_NS2; 1.
Pfam; PF01906; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01506; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                               Query Match
Best Local Similarity 84.3*
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q91RR6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DR
DR
DR
DR
DR
DR
MWW
KWW
SO
                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVIQMYTHVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRGDSR 123
                       79 IASPRGPVIQMYTNVDRDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                              19 TAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSTAAQIFLAICINGVCWIVYHGAGIRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369224; AAK54549.1;
Interpro; IPR004109; BCV_NS3.
Pfam: PF02907; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AA; 19131 MW; 8BD7FC2769DBD635 CRC64;
                                                                                                                                                                                                                                                                                           01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DRC-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86.3%; Score 891; DB 12;
96.1%; Pred. No. 8.2e-83;
tive 2; Mismatches 5;
                                                                                                                                                                                                                                                   181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Search completed: August 30, 2003, 19:18:19 Job time : 38.5921 secs
                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 96.1
Matches 171; Conservative
                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
```

```
nypothetical prote
transferrin-bindin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein
genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             polyprotein
polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - dour
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 polyprotein - marm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nitrate/nitrite se
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nitrate/nitrite se
probable exported
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         prote
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          angio-associated m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                glycosyl transfera
surface-associated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     proteinase (EC 3.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mosA protein - Rhi
probable periplasm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hypothetical prote
transferrin-bindin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             low density lipopr
                                                                                                                           (without alignments)
1168.492 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 chymotrypsin-like
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical
                                                                                                       August 30, 2003, 19:02:22 ; Search time 16.2134 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                               1032
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                   283308 seqs, 96168682 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                summaries
                                                                      - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRLJHD
B53308
B71284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNWVCJ
S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         E75392
F81196
JE0372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GNWVC3
S40770
GNWVCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GNWVTW
A45573
GNWVTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JC5620
JQ1303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AH1030
B71360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        139383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 138136
                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 r08839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              087603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S61441
                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                               US-09-965-594-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 915
1615
322
2508
264
323
333
398
716
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR_76:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8883.5
8857.5
8853.5
8853.5
8853.5
600.5
600.5
886.5
888.5
888.5
888.5
888.5
888.5
888.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
770.5
7
                                                                                                                                                                                                                                                   Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               76.5
76.5
76.5
76.5
                                                                                                                                                                                                 Perfect score:
                                                                       OM protein
                                                                                                                                                                                                                  Sequence:
                                                                                                                                                                                                                                                                                                      Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Database
                                                                                                         Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Title:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .
چ
```

probable aromatic	hypothetical prote	hypothetical prote		centromere protein	hypothetical prote	pyruvate dehydroge	ferredoxin-nitrite	transferrin-bindin	K9R protein - vari	transferrin precur	1rregular chiasm C	hypothetical prote	hypothetical prote	phosphodiesterase	aconitate hydratas
H83144	B83477	T04533	AB1775	S28261	T49806	DEHUPT	S16603	C81832	F72171	S33761	A49448	T35594	T16656	I53865	н81775
7	N	~	N	7	a	-	-	~	~	~	N	~	N	~	a
209	515	846	1334	2663	354	388	594	910	318	206	764	233	429	844	868
7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2	7.2	7.2
		'n	'n	ī.	75	75	75	75	74.5	74.5	5.4	74	74	74	74
75.5	75.	75.	75	75					-	-	-				

ALIGNMENTS

```
A; Molecule type: mRNA
A; Residues: 1-3011 <CHO>
A; Richan, S. W.; McCmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 1131-1141, 1992
A; Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship 1
A; Reference number: PQ0393; MUID:92268871; PMID:1316939
A; Accession: PQ0403
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A.Accessium: Accessive Acc
genome polyprotein - hepatitis C virus (strain HCV-1)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F;1616-1862/Próduct: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                 C; Species: hepatitis C virus
C; Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
C; Accession: A99166; PQ0403; PQ0404
R; Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Proco. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991
A; Title: Genetic organization and diversity of the hepatitis C virus.
A; Reference number: A39166; MUID:91172826; PMID:1848704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85.7%; Score 884.5; DB 1; illarity 84.3%; Pred. No. 7.1e-73; Conservative 9; Mismatches 14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E-b16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Experimental source: isolates
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA
A; Residues: 1577-1633 <CHA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Accession: A39166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Accession: PQ0404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    172;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 172
```

23

---LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST

3 KKGSVVIVGRIN----

à

~

```
genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
N;Cottains: capsid protein NS4s; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 83.3%
Matches 170; Conservative
                                                                                                                                                                                                                                                           Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Succession: Capsid protein C: envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4s; nonstructural protein C: envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS5 C; Species: hepatitis C vitus
C: Date: 19.49-1200 seaquence_revision 19-May-2000 *text_change 19-Jan-2001
C: Accession: $40770; PC1285
R: Okanoto, M; Accession: $40770; PC1285
R: Okanoto, M; Accession: $40770
A: Nolecule type: genomic RNA
A: Residues: 1.301 *CoKA
A: Residues: 1.501 *CoKA
A: Residues: 1.500 *CoKA
A: Re
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1005 RKGREILLGPADGMVSKGWRLLAPITAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1065 AAOTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGARSLTPCT 1124
                                                                                                          1125 CGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 ATQTFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
1005 RRGREILLGPADGMYSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                     CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.6%; Score 883.5; DB L.
llarity 84.3%; Pred. No. 8.8e-73;
Conservative 8; Mismatches 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1185 CTRCVAKAVDFIPVESLETIMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                        1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                              174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 172; Conserv
                                                                            54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                     q
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                ð
```

```
C:pecises: hepatitis C virus
A:Note: host Homo sapiens (man)
C:pecises: hepatitis C virus
A:Note: host Homo sapiens (man)
C:pate: 31-pec-1992 *Requence_revision 31-Dec-1992 *text_change 19-Jan-2001
C:Accession: A36814 A.M1546
R:Inchauspe, G: Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Submitted to Genamic structure of the human prototype strain H of hepatitis C virus
A:Reference number: A36814
A:Nolecule type: genomic RNA
A:Residues: 1-3011 A.TNC>
A:Residues: A8146; MUD:92052256; PMID:1658800
A:Title: Genomic structure of the human prototype strain H of hepatitis C virus: comp
A:Title: Genomic structure of the human prototype strain H of hepatitis C virus: comp
A:Reference number: A4146; MUD:92052256; PMID:1658800
A:Title: Genomic structure of the human prototype strain H of hepatitis C virus: comp
A:Reference number: A4146; MUD:92052256; PMID:1658800
A:Title: Genomic structure of the human prototope protein; glycoprotein; hydrolase; nonstruct
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome protein; glycoprotein; hydrolase; nonstruct
F:116-191/Product: andyor ervelope protein M status predicted ANED>
F:107-181/Product: andyor ervelope protein M status predicted ANED>
F:107-181/Product: andyor ervelope protein NS2 *status predicted ANED>
F:107-181/Product: nonstructural protein NS2 *status predicted ANED>
F:107-181/Product: nonstructural protein NS2 *status predicted ANED>
F:107-181/Product: nonstructural protein NS4 *status predicted ANED>
F:107-181
genome polyprotein - hepatitis C virus (strain H)
N/Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4a; nonstructural protein NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.1%; Score 878.5; DB 1,
83.3%; Pred. No. 2.5e-72;
tive 10; Mismatches 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174 CTRGVAKAVDFIPVESLETTMRSP 197
```

us-09-965-594-16.rpr

```
Query Match
Best Local Similarity 78.4%
Matches 160; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches 159; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: genomic RNA A; Residues: 1-3010 <TAK>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        C; Accession: A38465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                         ŝ
                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Necessian.

A; Experimental source: HCV-JT

A; Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)

C; Superfamily: hepatitis C virus genome polyprotein

C; Keywords: AFP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; S; 2-115/Product: capsid protein C *status predicted <EPN-
F; 116-191/Product: major envelope protein M *status predicted <ENN-
F; 190-729/Product: major envelope protein NSI *status predicted <NSI>F; 190-729/Product: monstructural protein NSI *status predicted <NSI>F; 100-106/Product: nonstructural protein NSI *status predicted <NSI>F; 1230-1237/Region: nucleotide-binding motif A (P-loop)

F; 1312-1317/Region: nucleotide-binding motif B
F; 1316-1319/Region: DEXH motif
                                                                        **Nolecule type: genomic RNA
A:Residues: 1-3010 <CHE>
CHE>
A:Residues: 1-3010 <CHE>
CHE>
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein; glycoprotein; hydrolase; nonstructura
C:Reywords: ATP; capsid protein C *status predicted <CHE>
F:116-191/Product: envelope protein W *status predicted <CHE>
F:130-139/Product: major envelope protein E *status predicted <NSI>
F:300-1006/Product: nonstructural protein NSI *status predicted <NSI>
F:1007-1615/Product: hepacivirin *status predicted <NSI>
F:1316-1317/Region: nucleotide-binding motif A (P-100P)
F:1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructut protein NS4s; nonstructural protein NS5 C;Species hepatitis C virus
C;Species: hepatitis C virus
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C;Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Virus Res. 23, 39-53, 1992
A;Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: A;Reference number: A45573; MUID:92295714; PMID:1318627
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping A;Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 867.5; DB 1
Pred. No. 2.6e-71;
18; Mismatches 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1185 CTRGVAKAVDFVPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84.1%;
79.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 163; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
```

```
R.TAKAMIZAWA, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, I. J. Virol. 65, 1105-1113, 1991
A.TILLE: Structure and organization of the hepatitis C virus genome isolated from hur A;Feference number: A38465; MUID:91140698; PMID:1847440
A;Accession: A38465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Across references: EMBL:MS8335; NID:q329770; PIDN:AAA72945.1; PID:q329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Sulfamily: hepatitis C virus genome polyprotein
C; Sulfamily: hydrolase; nonstructural protein M *status predicted <MEE>
C; 190.729/Product: nonstructural protein NSI *status predicted <NSI>
C; 1007-1615/Product: nonstructural protein NSI *status predicted <NSI>
C; 1007-1615/Product: hepacivirin *status predicted <NSI>
C; 11370-1006/Product: hepacivirin *status predicted <NSI>
C; 113712-1317/Region: nucleotide-binding motif B
C; 1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4 #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;206,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (EC 3.4.21.98) (nonstru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98
protein Ns4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 ATQIFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9; Gaps
                                                                                                                                                                                                                          Gaps
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                      6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3010;
                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1;
                                                                                                                                    83.1%; Score 857.5; DB 1; 78.4%; Pred. No. 2.2e-70; tive 20; Mismatches 15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82.7%; Score 853.5; DB 1
77.9%; Pred. No. 5.1e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 CTRGVAKAVDFIPVESLETTMRSP 197
```

a

```
A: Variety: isolate JKI
C; Date: 19-M97-2000 *text_change 23-Mar-2001
C; Accession: $18030; $33570; A48332; $18029
C; Caccession: $18030; $33570; A48332; $18029
C; Accession: $18030
S; Honda, M.: Raneko, S.: Massahi, U.: Robayashi, K.: Murakami, S.
submitted to the EMEL Data Library, September 1991
A. Poescription: A whole genome of hepatitis C virus cona was isolated from a single pa A. Reference number: $18038
A. Rocession: $18030
A. Mocelle type: genomic RNA
A. Residues: 1-3010 <4HONS
A. Cross-references: EMEL:X61596; NID:959478; PIDN:CAA43793.1; PID:959479
A. Cross-references: EMEL:X61596; NID:959478; PIDN:CAA43793.1; PID:959479
A. Cross-references: EMEL:X61596; NID:93119270; PMID:8380322
A. Title: Sequence analysis of putative structural regions of hepatitis C virus isolat A. Reference analysis of putative structural regions of hepatitis C virus isolat A. Reference analysis of putative structural regions of hepatitis C virus isolat A. Reference analysis of putative structural regions of hepatitis C virus isolat A. References: EMBL:X61591
A. Rocession: $33370
A. Moce: this sequence analysis of putative produce (NCEIN:121747, NCBIP:121748)
A. Note: this sequence is inconsistent with the nucleotide translation
A. Note: this acquence is inconsistent with the nucleotide binding; P-loop; polyprotein; A. Note: the authors translated the codon Acd for residue 43 as Pro, A. G. Superfamily: hepatitis C virus genome polyproted.
A. Note: the authors translated for codon Acd for residue 41 as Product: analysis of Superfamily: hepatitis C virus genome polyproted.
C. Superfamily: hepatitis C virus genome 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Cispecies: hepatitis C virus
Cibate: 19-May-2000 *text_change 19-Jan-2001
Cibate: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
Cibate: 10-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
Richamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina A;Reference number: JC5620; MUID:97366593; PMID:9223423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.4.21.98) (nonstru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Fil616-1862/Product: nonstructural protein NS4s *status predicted <N4A>
Fil863-2013/Product: nonstructural protein N84s *status predicted <N4B>
Fi2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
Fi2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
Fi196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  tch 81.6%; Score 842.5; DB 1; Length 3010; al Similarity 77.5%; Pred. No. 5.3e-69; 158; Conservative 20; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                               Cypecies: hepatitis C virus
Cypecies: hepatitis C virus
Cypecies: hepatitis C virus
Cypecies: hepatitis C virus
Cypecies: however, a virus
Cypecies: hepatitis C virus
Cypecies: hypotherizatis C virus
Cypecies: hypotherizatis M: Octauyama, Y: Nakagawa, M: Ohkoshi, S: Sugimura, T.; Shimot
Ryato, N: Hijikata M: Octauyama, Y: Nakagawa, M.; Ohkoshi, S: Sugimura, T.; Shimot
Ryato, N: Hijikata M: Octauyama, Y: Nakagawa, M.; Ohkoshi, S: Sugimura, T.; Shimot
Ryato, N: Hijikata M: Octauyama, Y: Nakagawa, M.; Ohkoshi, S: Sugimura, T.; Shimot
A:Title: Molecular choning of the human hepatitis C virus genome from Japanese patients
A:Reference number: Nag73; MID:9108850; PMID:2175903
A:Reference number: Nag73; MID:9221610; PIDN:BAA14233.1; PID:9221611
Ryato, N: Ohkoshi, S:: Shimotohno, K.
Proc. Jpn. Acad. 658, 219-223, 1980
A:Title: Japanese Isolates of the non-A, non-B hepatitis viral genome show sequence varia
A:Reference number: PS0085
A:Accession: PS0086
A:Access
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (isolate JKl)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructura
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATCINGVCWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   82.7%; Score 853.5; DB 1; Length 3010; 77.0%; Pred. No. 5.1e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23; Mismatches
                                                            1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 77.0
Matches 157; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
```

g

음 à

ò q õ

```
F;2018-3033/Product: nonstructural protein NS5 #status predicted <NO5>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                    .6%; Score 677; DB 1; .7%; Pred. No. 9.8e-54;
                                                                                                                                                                                                                                                26; Mismatches
                                                                                                                                               65.6%;
68.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA
A; Residues: 2678-2754 <CHA>
A; Cross-references: DDBJ:D10134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 65.4%
Best Local Similarity 69.3%
Matches 124; Conservative
                                                                                                                                                                                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1320-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A:Residues: 1-3033 <OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: mRNA
A; Residues: 2678-2729 <KAT>
                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                           Matches 123;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             139
                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Juliona polyprotein - hepatitis C virus (isolate HC-J6)
Wicontains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructus protein NS4s inonstructural protein NS4b; nonstructural protein NS5
C; pectes: hepatitis C virus
C; pate: 19-May-2000 #sequence_revision 19-May-2000 #tcxt_change 17-Nov-2000
C; Accession: J04303
C; pate: 19-May-2000 #sequence_revision 19-May-2000 #tcxt_change 17-Nov-2000
C; Accession: J04303
C; pate: 19-May-2000 #sequence_revision 19-May-2000 #tcxt_change 17-Nov-2000
C; Accession: J04303
C; pate: 19-May-2000 #sequence_revision 19-May-2000 #tcxt_change 17-Nov-2000
C; Accession: J04303
C; pate: 10-May-2000
C; p
                    A; Residues: 1-3014 <CHA>
A; Residues: 1-3014 <CHA>
A; Residues: 1-3014 <CHA>
A; Residues: 1-3014 <CHA>
A; Cross_references: GB:013184
A; Cross_references: GB:013184
A; Experimental source: genotype 5a, which predominates in South Africa
A; Note: the translation of the nucleotide sequence is not complete in this paper
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C ** status predicted <PE>
F; 192 - 389/Product: envelope protein N** ** status predicted <N** S: 384 + 408/Region: hypervariable ** status predicted <N** S: 389 + 730 + 730 + 700 + 730 + 750 + 730 + 750 + 730 + 750 + 730 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 750 + 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1006 RRGREIFLGPADDIKTSGWRLLAPITAYAQQTRGVLGAIVLSLTGRDKNEAEGEVQFLST 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQIFLATCINGVCWTVYHGAGIRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLIPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F:1620-1866/Product: nonstructural protein NS4a Estatus predicted <N4A>F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              73.7%; Score 760.5; DB 1; Length 3014; 69.6%; Pred. No. 1.9e-61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 69.6%;
Matches 142; Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
```

```
genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstreation Ns4a nonstructural protein NS4a; onstructural protein NS5
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: A40250; pp0397; pp0559
R; Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, Virology 188, 331-341, 1992
A; Title: Full-length sequence of a hepatitis C virus genome having poor homology to A; Reference number: A40250; MUID:92230232; PMID:1314459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCOSS-TEGERENCES: GB:D10562; GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
CS:Uperfamily: hepatitis C virus genome polyprotein
C:Reywords: ATP: capsid protein: envelope protein: glycoprotein; hydrolase; nonstruc
C:Reywords: ATP: capsid protein C fstatus predicted <CPC>
F:1-115/Product: capsid protein C fstatus predicted <CPC>
F:1-12-12-13/Product: envelope protein M fstatus predicted <MED>
F:192-139/Product: major envelope protein R fstatus predicted <MED>
F:390-733/Product: nonstructural protein NS1 fstatus predicted <NS1>
F:734-1010/Product: nonstructural protein NS2 fstatus predicted <NS2>
F:1011-1613/Product: hepatolytin fstatus predicted <NS2>
F:1214-1211/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A, Experimental source: isolate E-bl2
R; Kato, N.; Ootsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimot
Blochem. Blophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Réference number: PQ0554; MUID:92068204; PMID:1720309
A; Recession: PQ0559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           F;1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 *status predicted <NS5>
F;2018-303,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Cross-references: GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; PID:g221609 R;Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A;Reference number: PQ0393; WUID:92268871; PMID:1316939
    ö
                                                                                                                                       ö
                                                                                                                                                                                                                                                        79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRSP 197
                                                                            19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
.;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31; Indels
30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    65.4%; Score 675; DB 1;
69.3%; Pred. No. 1.5e-53;
Live 24; Mismatches 31;
```

us-09-965-594-16.rpr

9

Page

Oy 13 TATOTFLATCINGVCWTVYHGACTRIIASPKGPVTOMYTN 1 1 1 1 1 1 1 1 1	RESULT 14 B01104 nitrate/nitrite sensor protein (EC 2.7.3) NMB1249 [s C.Species: Neisseria meningitidis C.Species: Neisseria meningis R.Fettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffri Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Frist, H.; M.; Oin, H.; Vannathevan, J.; Gill, J.; Scarlato, V.; Gill, H.; Vannathevan, J.; Gill, J.; Scarlato, A; Mille: Complete genome sequence of Neisseria meningis A; Reference number: A81000; MUID:20175755; PMID:107103 A; Molecule type: DNA	A;Residues: 1-590 <pet> A;Cross-references: GB:AE002473; GB:AE002098; NID:9722 A;Experimental source: serogroup B, strain MC5B C;Genetics: A;Gene: NMB1249 C;Superfamily: nltrate/nltrite sensor protein narX C;Keywords: autophosphorylation: phosphohistidine; pho</pet>	Ouery Match 8.6%; Score 88.5; DB 2; Best Local Similarity 21.8%; Pred. No. 2.7; Matches 47; Conservative 28; Mismatches 76; Qy 28 EEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVC Db 213 EGGTPEFKQVGRCFNQMGGRLKILYDDLEGQVA	QY 88 QMTTNVDKDLVGWQAPQGSRSLTPCT	RESULT 15 C81911 nitrate/nitrite sensor protein (EC 2.7.3) NMA1418 (s C;Species: Neisseria meningitidis C;Species: 05-May-2000 #sequence_revision 05-May-2000 #te	R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.I.; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mur Nature 404, 502-506, 2000 A:Title: Complete DNA sequence of a serogroup A strair A; Reference number: A81775; MUID:2022556; PMID:107619 A:Status: preliminary
19 TAXAQOTRGEEGCOETSOTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78	RESULT 12 T08841 POLYPOTATION C; Species: douroucouli hepatitis GB virus A C; Species: douroucouli hepatitis GB virus A C; Species: douroucouli hepatitis GB virus A C; Species: 20-Sep-1999 *sequence_revision 20-Sep-1999 *text_change 17-Nov-2000 C; Accession: T08841 R; Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, ï.K. J; Gen. virol. 79, 41-45, 1998 A; Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A; Reference number: 216486; MUID:98120818; PMID:9460920 A; Accession: T08841 A; Status: translated from GB/EMBL/DDBJ A; Molecule type: mRNA A; Residues: 1-3005 <erk> A; Residues: 1-3005 <erk> A; Cross-references: EMBL:AF023425; NID:q2828599; PIDN:AAC40502.1; PID:g2828600 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: polyprotein</erk></erk>	<pre>Query Match Best Local Similarity 34.1%; Pred. No. 5.2e-15; Matches 56; Conservative 29; Mismatches 69: Indels 10: Gaps 3; Qy</pre>	QY 93 VDKDLVQWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKG 152 :	TESULT 13 T08839 polyprotein - marmoset hepatitis GB virus A C;Species: marmoset hepatitis GB virus A C;Species: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000 C;Accession: T08839 R;Erker, J.C.; Desal, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. Virol, 79, 41-45, 1998 A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A;Accession: T08839 A;Accession: T08839	7; PIDN otein	Ouery Match 24.3%: Score 251; DB 2; Length 2970; Best Local Similarity 27.8%; Pred. No. 1.8e-14; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Qy 3 KKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVS 52 :: :: :: :: Db 946 RRGDEVLIGVLNGVWELPPGFVPTAPVVVHHGRGFFGVVKTSMTGWDETEHVGNVVVLG 1005

```
C.M.: Moxon, E.R.; Rappuoli, R.;
gitidis serogroup B strain MC58.
0307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1.D.; Churcher, C.; Klee, S.R.; Motungall, K.; Quail, M.A.; Rajandre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           226488; PIDN: AAF41629.1; PID:9722
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            hosphoprotein; phosphotransferase
e) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [similarity] - Neisseria meningit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 in of Neisseria menigitidis 22491
1919
                                                                                                                                                                                                                                                    [similarity] - Neisseria meningit
                                                                                                                                                                                                                                                                                                                       ries, A.C.; Nelson, K.E.; Eisen,
Fleischmann, R.D.; Dougherty, B.
V.; Masignani, V.; Pizza, M.
                                                               ----PRPISYLKGSSGGPLLCP 161
| : :||| |:||
ELDLPAELCDFRGSSGSPILCD 1112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Indels 65; Gaps 7;
                       : i: : | |:: | ||
WSASDDVAVYPLPVGAKCLEPC 1055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CTCGSSDLYLVTRHAD---- 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | | ||:|:
|CLDGGSDVYVSIHHADCGTAAS 318
INVDKDLVGWQAPQGSRSLTPC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VCWIVYHGAGIRIIASPKGPVT 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             text_change 02-Feb-2001
                                                                                                                                                                                                                                                                                       text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 590;
                                                                                                                                                       | |
HEAPP 1154
                                                                                                                                   TMRSP 197
```

```
A; Molecule type: DNA
A; Residues: 1-590 <PAR>
A; Residues: 1-590 <PAR>
A; Residues: 1-590 <PAR>
A; Cross=**references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84658.1; PID:g738007
A; Experimental source: serogroup A, strain 22491
C; Genetics: MA1418
C; Genetics: MA1418
C; Superfamil: ntrate/nitrite sensor protein narx
C; Reywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase; s
F; 395/Active site: His (phosphohistidine intermediate) **status predicted*

**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status predicted*
**Active site: His (phosphohistidine intermediate) **status
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         88 QMYTNVDKDLVGWQAPQ----- 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       128 -----VIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAA----- 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            319 DLGKYHEEIFPIEYQNETLGRLLLSFPNGISLDEDDRILLQTLGYGLGVSLAGAKQEEEK 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28 EEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPVT 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match

8.6%; Score 88.5; DB 2; Length 590;
Best Local Similarity 21.8%; Pred. No. 2.7;
Matches 47; Conservative 28; Mismatches 76; Indels 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
```

Search completed: August 30, 2003, 19:20:29 Job time : 18:2134 secs

Run on:

```
ARITAGO Sequence 10331 Sequence 15 ARITATO Sequence 15 ARITATO Sequence ARITATO Sequence ARITATO Sequence ARITATO Sequence 108294 Sequence ARITATO Sequence ARITATO Sequence ARITATO Sequence ARITATO Sequence ARITATO Sequence BD08034 Hepatitis C 171894 Sequence 9 181885 Sequence 9 181885 Sequence 9 M62321 Hepatitis C AF387805 Synthetic AF387805 Synthetic AF387805 Synthetic AF387805 Synthetic AF387805 Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        D10749 Hepatitis C
AX395309 Sequence
AX454818 Sequence
AX127810 Sequence
BD081911 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF511948 Hepatitis
AR127809 Sequence
BD081910 Hepatitis
AR110828 Sequence
BD069982 Functiona
AR09606 Hepatitis
AR110831 Sequence
                                                                                                                                                                                                                                            AR179057 Sequence
AF268278 Pestiviru
AR11868 Sequence
106434 Sequence 48
109328 Sequence 8
                                                                                                                                                                                                                                                                                                           106440 Sequence 54
109329 Sequence 10
                                                                                                                                                                                                                                                                                                  AR118692 Sequence
                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    E08263 gRNA of
E08264 CDNA of
                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ALIGNMENTS
                                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HPCPLYPRE
AF387805
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AF387808
AF271632
AF387806
AF387807
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AF511948
AR127809
                                                                                                                                                                                                                                                                                                                                                       HPCPOLYP
                                                                                                                                                                                                                                                                                                                                                                                                                  108294
AR166930
AR301300
AR176483
BD080334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AX395309
AX454818
AR127810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AF009606
                                                                                                                                                                                                                                                                                                                                                                                             AR118723
BD091382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   E08263
E08264
HPCHCJ1
                                                                                                                                                                                                                                                                                                                                                                                     AR118722
                                                                                                                                                                                                                                                                                                                                                                AR118703
                                                                                                                                                                                                                                                                                                                                                                            AR118728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  171894
181885
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         E66293
        em_htg_hum:*
em_htg_inv:*
em_htg_other:*
                                                                                                                           em_htgo_other:*
                                                                                              em_sy:*
em_htgo_hum:*
                                                                                   em_htg_vrt:*
                                                    em_htg_pln:*
em_htg_rod:*
                                          em_htg_mus:*
                                                                        em_htg_mam:*
                                                                                                                  em_htgo_mus
                                                                                                                                                                                                             Query
Match Length DB
                                                                                                                                                                                                                                                                                                           6785
6785
7310
7310
7310
8316
8987
9185
9185
9185
                                                                                                                                                                                                                                                                                                                                                                                                                                                            9401
9401
9401
9401
9609
9609
9608
9608
9608
9608
9608
81505
8151
896.5
896.5
896.5
896.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 1
                                                                                                                                                                                                                                                                                      Result
                                                                                                                                                                                                                          .
8
                                                                                                                                                                                                                                                                                                                                       Command line parameters:
-WODEL-frame+_p2n.model -DEV=xlp
-WODEL-frame+_p2n.model -DEV=xlp
-WODEL-frame+_p2n.model -DEV=xlp
-Won2_1/USF70_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-Q3-/cqn2_1/USF70_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-DB-GenEmbl .OFWT=fastap -SUFPIX=rge -MINMARCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNTS=blts .START=1 -EMP=-
-UNTSER_LIGN=200 -THR_SCORE_pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE-LOCAL
-USFR-LSO9965594_eCGN_1_1.14686_erunat_29082003_151919_28310 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES_0 - WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XOAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELEXT=7
                                                                                (without alignments)
3147.423 Million cell updates/sec
                                                                                                                           1032
1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
                                                                       August 30, 2003, 19:18:33; Search time 2560.57 Seconds
      GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                   OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                        2888711 seqs, 20454813386 residues
                                                                                                                                                                                                                                             of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                               Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                  Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                 US-09-965-594-14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9b ba: *
9b_htg: *
9b_htg: *
9b_on: *
9b_ox: *
9b_pt: *
9b_vt: *
9b_vt: *
9b_vt: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_fun:*
em_hum:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_pat:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_in:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        еш_ош: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_ov:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_mu: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_pl:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        em_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GenEmbl:*
                                                                                                                                                            BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Perfect score:
Sequence:
                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                              rotal number
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Database :
                                                                                                                                                                                                                        Searched:
```

нер Нер

a

```
Location/Qualifiers
                                                                                                                                                                                                                                         386. .12508
/codon_start=1
 AF268278
AF268278.1 GI:9049956
                             type 1
                                                                                                                                                                                                                                  385
                                       Pestivirus
                             Pestivirus
                                                        Pestivirus
                                                                                                                                10864644
                                                                                                                                                                                                source
                                                                                                                                                                                                                                 5'UTR
CDS
                                      ORGANISM
                                                                                                                              PUBMED
 ACCESSION
          VERSION
KEYWORDS
SOURCE
                                                                 REFERENCE
                                                                          AUTHORS
                                                                                                                      MEDLINE
                                                                                                                                         REFERENCE
                                                                                                                                                  AUTHORS
                                                                                                                                                                    JOURNAL
                                                                                                             JOURNAL
                                                                                                                                                                                     FEATURES
                                                                                                                                                            TITLE
                                                                                           TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VRL 12-JUL-2000
           PAT 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113 GGTAGTGTTGTTATTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGCCGTAC 472
                                                                                                                                                                                                                                                                                                                    652
                                                                                                                                                                                                                                                                                                                                                                                                                                                       102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                          41
                                                                                                                                                                                                                                                                                                                                                                               GlySerValValIleValGlyArg1leAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                         AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                1 (bases 1 to 12734)
Hong, Z., Lai, W.C.H. and Lau, J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus
Patent: US 6326137-A 1 04-DEC-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
           linear
                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RNA
            DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AF268278 12734 bp Pestivirus type 1, complete genome.
      Sequence 1 from patent US 6326137.
                                                                                                                                                                                                                                                  (1-12734)
                                                                                                                                      /organism="unknown"
a 2604 c 3295 9
                                                                                                                                                                                                                                                  US-09-965-594-14 (1-197) x AR179057
                                     AR179057.1 GI:20220612
                                                                                                                                                                                   6.27e-65
930.50
94.87%
92.82%
                                                                          Unclassified
                                                                                                                                                                                                                 Similarity:
                                                       Unknown.
                                                                  Unknown
                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                          173
                                                                                                                                                                                                                                                                                                                                                               333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                142
                                                                                                                                                                                                                                                                     S
                                                                                                                                                                                                                                                                                                         55
                                                                                                                                                                                                                                                                                                                                             42
                                                                                                                                                                                                                                                                                                                                                                                62
                                                                                                                                                                                                                                                                                                                                                                                                  593
                                                                                                                                                                                                                                                                                                                                                                                                                     82
                                                                                                                                                                                                                                                                                                                                                                                                                                      553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          713
                                                                                                                                                                                                                        Query Match:
                                                                                                                                 source
                  DEFINITION
ACESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
                                                                                                                                                  BASE COUNT
                                                                                                                                                                                                                 Best Local
                                                                 ORGANISM
                                                                                                                                                                                       Pred. No.:
                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 2
AF268278
                                                                                   REFERENCE
                                                                                            AUTHORS
                                                                                                                      FEATURES
AR179057
                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         合
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ő
                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  음
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
OKNIGEDLHNOIGGWATCVPEDOLLYKGGSIESCKWGGYOFKESEGEPHYPTGKKL
ENETGYRLVDSTSCNREGVAIVPQGTLKCKIGKTTVQVIAMDTKLGPWPCRPYEIISS
EGPVEKTAGTRYTKTIKNKYFEPRDSYFQQYMLKGEYQYWFDLLEVDHHNDYFAESI
LVVVVALLGGTRYVLMILYTWYLSEQKALGIQYGSGEVVAMGNLLTHNNIEVYTYELL
LYLLIREESVKRWYLLLHTILVYHPIKSYIYILLMIGDYVKADSGGGEYLGKIDLCY
TVVLIVIGLIIARRDPTIVPLVTIMAALRYTELTHQPGVDIAVAWTITLLMVSYYDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KNCTPACLPRUTKIVGPGRFDTNAEDGKILHEMGGHLSEVLLLSLVVLSDFAPETASV
MYLILHFSIPOSHVDVMDCDKTOLNLTVELTTADVIPGSVWNLGRWVCIRPNWMPYET
TVYLAFBESGOVKLVLRALRDLTRINNAATTTAFLICLYTYRGGWVGSILMLLLIT
GVOGHLDCKPEFSYAIARDENIGOLGAGGLTTTWFESPGMKLEDTWVIAMCEDGKLM
YLQRCTRETYKLAILHTRALPYSVVFKKLFPGRKOEDVVEMINNFEPGLCPCDAKPIV
RGKFNTTLLNGPAFQMVCPIGWTGTVSCTSFNMDTLATTVVRTYRRSKPFPHRQGCIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LTLDPMYYMERKVIEEISGGTNIISRLVÄALIELMWSMEEEESKGLKKFYLLSGRLRN
LIIKHKVRNETVASMYGEEEVYGMPKIMTIIKASTLSKSRHCIICTVCEGREWKGGTC
PKCGRHGKPITCGMSLADFEERHYKRIFIREGNFEGMCSRCGGKHRRFEMDREPKSAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TGMAYTHQGGISSVDHVTAGKDLLVCDSMGRTRVVCQSNNRLTDETEYGVKTDSGCPD
GRACYYLDREAVNIGSSKGAVYHLQKTGGEFYCVTASGTPERPDLKNIGGWSGLDIFE
ASSGRVVGRVKVGRYKSGRRAVINGTHASGIQTVSRNTADLTEAVNKITSNNRGDFKQITLA
TGAGKTTELPRAVIEEIGRHKRVLVLIPLRAAAESVYQYMRLKHPSISFNLRIGDMKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LORHEWNKHGWCNWYNIEPWILVMNRTQANLTEGOPPRECAVTCRYDRASDLNVVTGA
RDSPT PLTGCKKGKNFSFAGILMRGPCNFEIAASDVLFKEHERISMFQDTTLYIVDGL
TNSLEGARGGTAKLTTWLGKOLGILGKKLENKSKTWFGAYAASPYCDVDRKIGYIWYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YFRYKWIQCILSLVSGVFLIRSLIYLGRIEMPEVTIPWRPLTLILLYLISTTYTR
WKVDVAGLLLQCVPILLLVTTLWADFLTLILILPTYELVKLYYLKTVRTDIERSWLGG
IDYTRVDSIYDVDESGEGVYLFPSRQKAQGNFSILLPLIKATLISCVSSKWQLIYMSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YCAECNRLHPAEBGDEWAESSMLGLKITYFALMDGKVYDITEWAGCQRVGISPDTHRV
PCHISFGSRWPFRQEYNGFVQYTARGQLFLRNLPVLATKVKMLMVGNLGEEIGNLEHL
GWILRGPAVCKKITEHEKCHINILDKLTAFFGIMPRGTTPRAPVRFPTSLLKVRRGLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GDMATGITYASYGYFCQMPQPKLRAANVEYSYIFLDEYHCATPEQLAIIGKIHRFSES
IRVVAMTATPAGSVTTTGQKHPIEEFIAPEVMKGEDLGSQFLDIAGLKIPVDENKGNM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LPDLDTVIDTGLKCEKRVRVSSKIPFIVTGLKRMAVTVGEQAQRRGRVGRVKPGRYYR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SQETATGSKDYHYDLLQAQRYGIEDGINVTKSFREMNYDWSLYEEDSLLITQLEILNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LLISEDLPAAVRNIMARTDHPEPIQLAYNSYEVQVPVLFPKIRNGEVTDTYENYSFLN
ARKLGEDVPVYIYATEDEDLAVDLLGLDWPDPCNQOVVETGRALKQVTGLSSAENALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 (bases 1 to 12734)
Lai.V.C.H. and Hong, Z.
Direct Submission
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research
Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GKTKSKNTQDGLYHNKNKPQESRKKLEKALLAWA11AIVLFQVTMGENITQWNLQDNG
TEGIQRAMFQRGVNRSLHGIWPEKICTGVPSHLATDIELKTIHGMMDASEKTNYTCCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVD
QDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHANVIPVRRKGDSRGSLLSPRPISYLKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LVFVPTRNMAVEVAKKLKAKGYNSGYYYSGEDPANLRVVTSQSPYVIVATNAIESGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /translation-"MELNTNEGSGSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SYSDTREEGATKKKTQKPDRLERGKWKIVPKESEKDSKTKPPDATIVVEGVKYQVRKM
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                 Lai, V.C., Zhong M., Skelton, A., Ingravallo, P., Vassilev, V., Donis, R.O., Hong, Z. and Lau, J.Y.
Generation and characterization of a hepatitis C virus NS3 J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /organism="Pestivirus type 1"
/mol_type="genomic RNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /product-"polyprotein"
/protein_id-"AAF82566.1"
/db_xref-"G1:9049957"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /mol_type="genomic RN2/db_xref="taxon:11099"
```

```
LIGYOLI TRIESOY TELETTI TATES INCORDILIZED SAMMENT INCORDING THE CONTROLLED TO SAMMENT INCORDING THE CONTROLLED SERVICE THE CONTROLLED SOURT TO STRUCK TO STRUCK TO STRUCK THE CONTROLLED SOURT THE CONTROLLED 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TPLFNIFDKVRREWDSFNEPVAVSFDTRAWDTQVTSKDLQLIGEIQKTYYKKEWHKFI
TTDHMTEVPVTTANCEVYTRAGGNSQPDTSAGNSMLWTTMMYRECESTCVPYK
SFNRVARIHVGENDEFLITEKGLGLKFANKGWILHEAGKPOKITEGEKMKVAYRFED
IEFCSHTPVPVRWSDNTSSHMAGRDTAVILSKMATRLDSSGERGTTAYERAVAFSFLL
MYSWPLVRYRICLLVLSGOPEPDPSKHATYYYKGDFGAKVOVIGRUESELKRTGFEK
LAMLANLSLSTLGINTHTRKR I 1000VAIGREGNWLVNADRLISSKTGHLY PPKGF
TLQGKHYEQLQLRTETNPVMGVGTERYKLGPIVNLLRRLKILLMTAVGVSS
GDVEKIMGAISDYAAGGLEFVKSQAEKIKTAPLFKENAEAAKGYVQKFIDSLIENKEE
IRYGIMGTHTALYKSIAARLGHETAPATLVIKMAFGGESVSHVKORAVDLVYYYY
MINESFPGDSETQOGRRYVAELFGALATTYYKMAHNLSKVVEPALAYLFYATSA
LKMFTPTRLESVYILSTTTYKTYLSIRKGKSDGLLGTGISAAMSIISQNEVSVOISVM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGTAGTGTTGTTATTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12734
178
5
9
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3293 g 2802 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-14 (1-197) x AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2.55e-63
910.50
93.85%
91.28%
88.23%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2608 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3'UTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ç
```

```
1046
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   173
            AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                      93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                         ::::::||| :::::||| 867 CGGAGGGGCCGGAGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         74 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       94 AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                       ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                   5360
175
9
11
11
                                                                                                                                                                                                                                                                                                                                                                                             3 LysLysGlySerValValIleValGlyArgIleAsn------
                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                   1 (bases 1 to 5360)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 53 21-NOV-2000;
Location/Qualifiers
1. 5360
                                                                                                     5360 bp us 6150087.
                                                                                                                                                                                                                                                            1145
                                                                                                                                                                                                                                                                                                                                            Indels:
                                                                                                                                                                                                                                                                                                                                                                        (1-5360)
                                                                                                                                                                                                                                              /organism="unknown"
1623 c 1532 a
                                                                                                    AR118686
Seguence 53 from patent AR118686
AR118686.1 GI:14100596
                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-14 (1-197) x AR118686
                                                                                                                                                                                                                                                                                                 5.32e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                          Unclassified.
                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                          ĸ
                                                                                                                                                                                                                                                           1060
                                                                                                                                                       Unknown
                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1287
                                         182
162
                                                                                                                                                                                                                                                                                                                                            Query Match:
                                                                                                                         ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                     ..
Q
                                                                                                                DEFINITION
                                                                                                                                                                                                                                                           BASE COUNT
                                                                                                                                                                                                       TITLE
JOURNAL
FEATURES
                                                                                  RESULT 3
AR118686
                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                      ORIGIN
                                                                                                                                                                                                                                                                                                   Pred
                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                             ò
                  a
                                         ò
                                                           g
```

	RESCUT 5 109328	Alignment Scores: 5.32e-63 Length: 5360 Pred. No.: 901.50 Marches: 175 Score: 90.204 Conservative: 9 Best Local Similarity: 85.784 Mismatches: 11 Query Match: 6.354 Indels: 9 DB: 6.354 Indels: 1 US-09-965-594-14 (1-197) x 109328 (1-5360)	0y 3 LysLysGlySerValValIleValGlyArglleAsn	Oy 15LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33	QY 34 ThrSerGinThrGlyArgAspLysAsnGinValGluGlyGluValGlnIleValSerThr 53	Oy 54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73	Oy 74 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal 93	Oy 94 AsplysAspleuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 113	Qy 114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133	0y 134 ArgGlyAspSerArgGlySerLeuteuSerProArgPro11eSerTyrLeuLysGlySer 153	Qy 154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
Db 1347 TCGGGGGGTCCCCTGTTGTGCCCCGGGGGCCCGTGGGCATATTAGGGCCGCGGTG 1406 Qy 174 CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193 1407 TGCACCCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGAACC 1466 Qy 194 MetArgSerPro 197	106434 106434 5360 bp DNA 11near PAT 02-DEC-1994 DEFINITION Sequence 48 from Patent EP 0318216. ACCESSION 106434 161:590311 ACCESSION 106434 161:590311 REYMORDS 106434 161:590311 REYMORDS 106434 161:590311 CRGANISM 106434 161:590311 CRGANISM 106434 161:59031 CRGANISM 106434 161:59031 CRGANISM 106434 165360 CRGANISM 106436 165360 CRGANISM	Alignment Scores: 5.32e-63 Length: 5360 Pred. No.: 901.50 Matches: 175 Scorer 1 175 Bercent Similarity: 90.20% Conservative: 9 Best Local Similarity: 85.78% Mismatches: 11 Ouery Match: 6.35% Indels: 9 DB: 6.35% 106434 (1-5360)	Qy 3 LysLysGlySerValValIleValGlyArglleAsn	Qy 15LeuSerGlyAspThrAlaTyralaGlnGlnThrArgGlyGluGluGlyGySGlnGlu 33	Oy 34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 53 111111 1111111111111111111111111	Oy 54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73 [11111111111111111111111111111111111	Oy 74 AlaGlyThrArgThrILeAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal 93 	Oy 94 AsplysAspleuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 113	Oy 114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133 	<pre>Qy 134 ArgdlyAspSerArgGlySerLeubeuSerProArgProIleSerTyrLeuLysGlySer 153</pre>	Oy 154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173

Qy 154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173	NESULT 7 106440	ignment Scores: 6.86e-63 ore: 901.50 ccent Similarity: 90.20% st Local Similarity: 85.78% ery Match: 6	0y -965-994-14 (1-197) x 106440 (1-6785) 0y 3 LysLysGlySerValValIleValGlyArgIleAsn
		Alignment Scores: Alignment Scores: Score: Score: Score: 901.50 Matches: Percent Similarity: 90.20% Conservative: Mismatches: Query Match: 87.35% Indels: DB: 0S-09-965-594-14 (1-197) x AR118692 (1-6785)	Db 1140 GGCGGGGGGGGGGGGGCGGGTGGATGGATGGTCTCCAAGGGTGGAGG 1199 Qy 15

9

Page

Oy 154 SerGl Db 1620 TCGGC Qy 174 CySTP Db 1680 TGCMC Qy 194 Metha	SULT 9 COSSION COSSION CESSION CESSION TWORDS URCE ORGANIS AUTHORS AUTHORS SOUR	Alignment Scores: Pred. No.: Score: Fercent Similarity Best Local Similar Ouery Match:	0y 3 Lyst, 0y 3 Lyst, 0y 3 Lyst, 0y 1665 cGcn; 15 Ly 17 17 17 17 17 17 17 17 17 17 17 17 17
0y 154 SerGlyGlyProLeuLcuCysProAlaGlyHisAlaValGlyIlePheArgalaAlaVal 173	SULT 8 5029 CUS CUS CUS CUS CUS CUS CUS CU	Alignment Scores: Pred. No.: Pred. No.: 901.50 Matches: Pacore: Procone: Procone: Percent Similarity: 90.208 Mismatches: 175 Conservative: 9 Best Local Similarity: 87.358 Indels: 1 US-09-965-594-14 (1-197) x 109329 (1-6785)	0y 3 LysLysGlySerValValIleValGlyArgIleAsn

```
:::|||
AGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1724
                                                                                                                                                                                                                                                                                                                                 PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlyaspSerargGlySerLeuLeuSerProargProlleSerTyrLeuLysGlySer 153
                                                                                                                                                                                                                                                                                                      LysGlySerValValIleValGlyArgIleAsn-----
                                                                                                   linear
                                                                                                                                                                                                                                      7310
175
9
11
                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                          .nown.
classified.
(bases 1 to 7310)
thien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 74 21-NOV-2000;
1..7310
1..7310
T.7310
                                                                                                   DNA
                                                                                                  696 7310 bp 1
nce 74 from patent US 6150087.
696.1 GI:14100606
                                                                                                                                                                                                                                                                                         (1-197) x AR118696 (1-7310)
                                                                                                                                                                                                                                      7.44e-63
901.50
90.20%
85.78%
87.35%
                                                       y:
rity:
                                                                                                                                       5
```

```
.>7310
                                                                                                                    11111111111
2265 ATGAGGTCCCCG 2276
                                                                                                       194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                    plasma.
                    154
                                                                                                                                                                                                                                                                                                                                                                                        source
                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                              AUTHORS
TITLE
                                                                                                                                                                                          ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                    JOURNAL
COMMENT
                                                                                                                                                 RESULT 11
HPCPOLYP
                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                  SOS
                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                          SOURCE
                                        g
                                                              ò
                                                                                 g
                                                                                                        ö
                                                                                                                            a
                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1905 GCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTA 1964
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     965 GACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACT 2024
                                                            1665 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1724
                               PAT 02-DEC-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer 153
2085 CGGGGTGATAGCAGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCC 2144
                     SerGlyGlyProLeuLcuCysProAlaGlyHisAlaValGlyIlePhcArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyGlnGlu
                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               LysLysGlySerValVallleValGlyArglleAsn-
                                                                                                                                                                      DNA
                                                                                                                                                                                                                                                   1 (bases 1 to 7310)
Houghton, M., Choo, Q. K. and Kuo, G.
Patent: WO 8904669-A 15 01-JUN-1989;
Location/Qualifiers
1. 7310
                                                                                                                                                                 Sequence 15 from Patent WO 8904669.
                                                                                                                                                                                                                                                                                                                                                                                                             Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-14 (1-197) x I09331 (1-7310)
                                                                                                                                                                                                                                                                                                             /organism="unknown"
2218 c 2058 g
                                                                                                                                                                                                                                                                                                                                                                 7.44e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                   2265 AIGAGGICCCCG 2276
                                                                                                                                                                                                     109331.1 GI:587966
                                                                                                        194 MetArgSerPro 197
                                                                                                                                                                                                                                                Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                         ิส
                                                                                                                                                                                                                                                                                                                         1495
                                                                                                                                                                                                                          Unknown.
                                                                                                                                                                                                                                       Unknown
                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ~
                     154
                                         2145
                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                              DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                      source
                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                             I09331
                                                                                                                                                                                                                                                                                                                                                                              Score:
                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                         a
                                                                õ
                                                                                                        ò
                      ò
```

```
ADARYCSCIMMALLISOARALEMINATURASIAGHUSEFUFECERMYLKOKWP

ADARYCSCIMMALLISOARALEMINATURASIAGHUSEFUFECERMYLKOKWP

GAVYTFYGWPLLILLIALAPORAYALDTEVAASACGOVUVGLAALTGSPYKRYISW

LUMALQYFLTRVEAQLHWIPPLANGGRADAYILLACAVHPTLVFDITKLLAVFGPLW

ILQASLLKVPFKROGLRECALARKKIGGHVQMYIKIACANTGTYYWHLTPLRD

MAHNGIRDLAVAVEWYSOMETKIITWGADTAACGIINGLPVSARRGREILLGPAD

GWYSKGWRLLAPITAYAQOTRGLIGGIITSTJGRDKNOVEGEWOIVSTAAQTFLATCI

NGVCWTVHGAGTRTIALSPREPUSPROVIOWTINVODLYGRPAPOGESBLIPCT

VTRHADVIPVRGATRIALSPREPUSPROVIOWTINVODLYGRPAPOGESSILY

VTRHADVIPVRATIGPEGANGSAHGIDPNIKTGVYRITTGSPTTYSTVGKTLADGGGS

GGAYDIICODECHSTDATSILGIGTVLDQAETAGARAYITRGSSTTYPPGGSTRADGGGS

GGAYDIICODECHSTDATSILGIGTVLDQAETAGARAYITRGSPTTYSTVGKTLADGGGS

VALSTTGEIPPFGRAIPEVIRGRAHGIDPNIKTGVYRITTGSPTTYSTVGKTLADGGGS

VSTTTGEIPPFGRAIPLEVIRGRAHGIDPNIKTGVTTOTVATATATPGGSTYPRIED

VSVIPTSGDVVVVATDALLATTGTGTTATGSPSTYDCTTOTVOTVDFSLIDPTFTITLDQD

VSVIPTSGDVVVVATDALLATTGTGTTATGSPSTYDCTTOTVOTVDFSLIDPTFTITLDQD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AQAPPPSWDQMWKCLIRLKPTLHGPTPLLYRLGAVQNEITLTHPVTKYIMTCMSADLE
VVTSTWVLVGGVLAALAAYCLSTGCVVIVGRVVLSGKPAIIPDREVLYREFDEWEECS
QHLPYIEQGMMLAEQFKQKALGLLQTASRQAEVIAPAVQTNWQKLETFWAKHMWNFIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GIVPAKSVCGPVYCFTPSPVVVGTTDRSGAPTYSWGENDTDVFVLMNTRPPLGNWFGC
TWNNSTGFTRYVCGAPRCVIGGAGNNTALGPTDCFRRHPDATYSRCGSGDWITPRCLVD
YPRLWHYPCTINTIFKIRMYVGGYEHRLEAACNWTRGERCDLEDRDRSELSPLLIT
TTOWQVLPCSFTTLPALSTGLIHLHQNIVDVQYLYGYGSSIASWAIKWEYVVLLFLLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TAILSSLTVTQLLRRLHOWISSECTTPCSGSWLRDIWDWICEVLSDFKTWLKAKLMPQ
LPGIPFVSCQRGYKGVWRVDGIMHTRCHCGAEITGHVKNGTWRIVGPRTCRNMWSGTF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LRAYMNTPGLPVCQDHLEFWEGVFTGLTHIDAHFLSQTKQSGENLPYLVAYQATVCAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GIQYLAGLSTLPGNPAIASLMAFTAAVTSPLTTSQTLLFNILGGWVAAQLAAPGAATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FVGAGLAGAAIGSVGLGKVLIDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VRL 02-AUG-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /translation="GCPERLASCRPLTDFDQGWGPISYANGSGPDQRPYCWHYPPKPC
2085 CGGGGTGATAGCAGGGCAGCCTGCTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCC 2144
                                                                                                                                                                                                                                    SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Original source text: Hepatitis C virus, cDNA to viral RNA, clones K9-1 through 15e, isolated from chimpanzee (individual 910) blood
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             braft entry and printed sequence for [1] kindly submitted by
M.Houghton, 22-FEB-1990. Chiron Corporation, 4560 Horton Street,
                                                                                                                        1145 TCGGGGGGTCCGCTGTTGTTGTTGTTGTTGTTGTTTAGGGCCGGGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (bases 1 to 7310)
Choo, 0.-L., Richman, K. and Han, J.
The nucleotide sequence of the Hepatitis C viral genome
Unpublished (1990)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HPCPOLYP 7310 bp ss-RNA linear Hepatitis C virus polyprotein gene, partial cds. M32084.1 GI:329875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       polyprotein.
Hepatitis C virus
Hepatitis C virus
Viruses: SSRNA positive strand viruses,
Hepatilvius.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1. 7310
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/db_xref="taxon:11103"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /protein_id="AAA45677.1"
/db_xref="GI:329876"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      M. Houghton, 22-FEB-15...
Emeryville CA 94608
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /note="polyprotein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /codon_start=3
```

```
REFERENCE
AUTHORS
AR118703
                                                                                                                                                                                                                                                                                                                      ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                   Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Óγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
         PEFFEELDOYRELING THE TOTAL TOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1904
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PINAYTTGPCTPLPAPNYTFALWRVSAEEYVEIRQVGDFHYVTGMTTDNLKCPCOVPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1665 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaGlyThrargThrileAlaSerProLysGlyProVall1eGlnMetTyrThrAsnVal 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                      7310
175
9
11
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LysLysGlySerValVallleValGlyArglleAsn-----
                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-14 (1-197) x HPCPOLYP (1-7310)
                                                                                                                                                                                                                                                                                   2058 g
                                                                                                                                                                                                                                                               SIEPLDLPPIIQRL
2218 c 20
                                                                                                                                                                                                                                                                                                                                                      7.44e-63
901.50
90.20%
85.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1111111111
2265 ATGAGGTCCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                   ď
                                                                                                                                                                                                                                                                                 1495
                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2025
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1725
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    46
                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                                                53
                                                                                                                                                                                                                                                                                                                                                                                                                                          73
                                                                                                                                                                                                                                                                                                                                                     33
                                                                                                                                                                                                                                                                                                                                                               ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TCGGGGGTCCGCTGTTGTGCCCCGCGGGCACGCCGTGGGCCATATTTAGGGCCGCGGTG
                                                                                                                                                                                                                                                                                                                                                    ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
   linear
                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
   ONA
                                                                                                                                                                                                                                                                                                           LysLysGlySerValVallleValGlyArgIleAsn
                                                                                   Chien, D.Y.

Chien, D.Y.

NANDY diagnostics and vaccines
Patent: US 6150087-A 88 21-NOV-2000;
Location/Qualifiers
1. 8316
 8316 bp 1
us 6150087.
                                                                                                                                                                                                                                                                                     (1-8316)
                                                                                                                                                     /organism="unknown"
2529 c 2345 g
                                                                                                                                                               2345 g
            88 from patent
                                                                                                                                                                                                                                                                                     US-09-965-594-14 (1-197) x AR118703
                                 AR118703.1 GI:14100613
                                                                                                                                                                                                         8.55e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATGAGGTCCCCG 3282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MetArgSerPro 197
                                                                           Unclassified
                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
AR118703
Sequence 8
AR118703
                                                                                                                                                                 1671 a
                                                       Unknown
                                                                 Unkaown
                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                   2791
                                                                                                                                                                                                                                                                                                                                                                                                                                                               2851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3091
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3271
                                                                                                                                                                                                                                                                                                         m
                                                                                                                                                                                                                                                                                                                                                   15
                                                                                                                                                                                                                                                                                                                                                                                              34
                                                                                                                                                                                                                                                                                                                                                                                                                                         54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174
                                                                                                                                                                                                                                                      Query Match:
          DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                            source
                                                                                                                                                                                                             ..
0N
                                                                                                          TITLE
JOURNAL
FEATURES
                                                                                                                                                                 BASE COUNT
```

σ

```
114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallieProValArgArg 133
114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallieProValArgArg 133
3692 TGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCCGG 3751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                         54 AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                               SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                linear
                                                                                                                                                                                                                                                                                             3 LysLysGlySerValValIleValGlyArgIleAsn------
                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                   1 (bases 1 to 9185)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 122 21-NOV-2000;
Location/Qualifiers
1, 9185
                          9185 bp E Sequence 122 from patent US 6150087. ARI18722 ARI18722.1 GI:14100632
                                                                                                                                                                      1938
                                                                                                                                                                                                                                                                          US-09-965-594-14 (1-197) x AR118722 (1-9185)
                                                                                                                                                            /organism="unknown"
2790 c 2608 g
                                                                                                                                                                    2608 g
                                                                                                                                                                                                         9.52e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MetArgSerPro 197
                                                                                              Unclassified
                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                      1849 a
                                                                           Unknown
                                                                                      Unknown
                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           194
                                                                                                                                                                                                                                               Query Match:
                                                                                                                                                      source
                                      DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                           Pred. No.:
                                                                                                              AUTHORS
TITLE
JOURNAL
          RESULT 14
AR118722
LOCUS
                                                                                                      REFERENCE
                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                       οχ
                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                             οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                      3072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3372
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3252
                                                                                                                                                                                                                                                                                                                                                                                                                                  3433 CGGGGTGATAGCAGGGGCAGCTGCTGTCGCCCGGCCCATTTCCTACTTGAAAGGCTCC 3492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                 73
                                                                                                                                                                                                                                                                                                     3013 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                          3193 GCTGCCCAAACCTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCACGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                  3 LysLysGlySerValValIleValGlyArgIleAsn----
                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
                                                                                   Unclassified.
1 (bases 1 to 8987)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-A 137 21-NOV-2000;
Location/Qualifiers
1. 8987
                            Sequence 137 from patent US 6150087.
AR118728
AR118728.1 GI:14100638
                                                                                                                                                                                                          Matches:
                                                                                                                                                                                                 Length:
                                                                                                                                                                                                                                                                US-09-965-594-14 (1-197) x AR118728 (1-8987)
                                                                                                                                               /organism-"unknown"
2735 c 2547 a
                                                                                                                                                                                              9.3e-63
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11111111111
3613 ATGAGGTCCCCG 3624
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MetArgSerPro 197
                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                             N
                    AR118728
                                                                                                                                                            1807
                                                                  Unknown
                                                                           Unknown
                                                                                                                                                                                        Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194
                                                                                                                                                                                                                                     Query Match:
                          DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                           source
                                                                          ORGANISM
                                                                                                     AUTHORS
TITLE
JOURNAL
FEATURES
                                                                                                                                                           BASE COUNT
RESULT 13
AR118728
LOCUS
                                                                                             REFERENCE
                                                                                                                                                                     ORIGIN
                                                                                                                                                                                                          Score:
                                                                                                                                                                                                                                                                                                     đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ογ
                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5794 TIGCIGGCGCCCAICACGCGCTACCAGCACAAGGGCCCTCCTAGGGTGCAIAATC 5735
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5374 resesses receserations and recesses recesse
                                                                                                         PAT 16-MAY-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               34 ThrSerGlnThrGlyArgAspLysAsnGLnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 LysLysGlySerValVallleValGlyArglleAsn-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                           DNA
                                                                                                                                                                                                                                                                                                                                                                                  21-NOV-2000;
                                                                                                   AR118723 9185 bp C
Sequence 123 from patent US 6150087.
AR118723.1 GI:14100633
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1849
                                                                                                                                                                                                                                                                                                      1 (bases 1 to 9185)
Chien, D.Y.
NANBV diagnostics and vaccines
Patent: US 6150087-a 123 21-NOV-
Location/Qualifiers
e 1. 9185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-14 (1-197) x AR118723 (1-9185)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="unknown"
1 2608 c 2790 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2790 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9.52e-63
901.50
90.20%
85.78%
87.35%
3932 ATGAGGTCCCCG 3943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MetArgSerPro 197
                                                                                                                                                                                                                                                                                      Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                    Unknown.
                                                                                                                                                                                                                                                         Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15
                                                 RESULT 15
AR118723/c
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                臽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

```
August 30, 2003, 19:13:57; Search time 182.939 Seconds
OM protein - nucleic search, using frame_plus_p2n model
                                                                  Run on:
```

(without alignments) 2906.924 Million cell updates/sec 1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197 US-09-965-594-14 1032 Perfect score: Sequence: Scoring table:

0.5 Xgapop 10.0 , Xgapext Ygapop 10.0 , Ygapext Fgapop 6.0 , Fgapext Delop 6.0 , Delext **BLOSUM62**

2552756 seqs, 1349719017 residues Searched:

Total number of hits satisfying chosen parameters: Minimum DB seq length: 0 Maximum DB seq length: 200000000

5105512

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:

-MODEL_frame+_p2n.model -DEV-xlp
-O-/Cogn_2_1/GSPTO_app_0001/OSO9965594/runat_29082003_151918_28302/app_query.fasta_1.2872
-O-/Cogn_2_1/GSPTO_app_001/OSO9965594/runat_29082003_151918_28302/app_query.fasta_1.2872
-DB-N_Geneseq_19Jun03 -OFMT-fastap -SUFFIX-=ng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST-45 -DOCALIGN=200 -THR_SCORE-pct -THR_MXX-100 -THR_MIN-0 -ALIGN=15
-MODE-LOCAL -OUTFWT-ptc -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-2000000000
-USER-USO9965594_GCGN_1_1-1412_ernat_29082003_151918_28302 -NCPU-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEXT=7

/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*/SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:* /SIDSI/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*/SIDSI/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*/SIDSI/gcgdata/geneseqn-embl/NA1994.DAT:*/SIDSI/gcgdata/geneseqn-embl/NA1994.DAT:*/SIDSI/gcgdata/geneseqn-embl/NA1995.DAT:*/SIDSI/gcgdata/geneseqn-embl/N /SIDS1/gogdata/geneseg/genesegn-embl/NA1996.DAT:*/SIDS1/gogdata/geneseg/genesegn-embl/NA1997.DAT:*/SIDS1/gogdata/geneseg/genesegn-embl/NA1998.DAT:* /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:* /SIDS1/gcgdata/geneseg/geneseqn-emb1/NA2000.DAT:* N_Geneseq_19Jun03:*

| SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
| SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1983.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1984.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1984.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1986.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1986.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1986.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1986.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1986.DAT:*
| SIDS1/gcgdata/geneseq-geneseqn-embl/NA1986.DAT:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

/SIDS1/gcgdata/geneseg/genesegn-embl/NA2002.DAT:*/SIDS1/gcgdata/geneseg/genesegn-embl/NA2003.DAT:*

and is derived by analysis of the total score distribution.

SUMMARIES Description		10 Hepatitis C v	Hepatitis C v	9 Hepatitis C v	Hepatitis C v	Hepatitis C v.	Hepatitis C v.	Mepatitis C V	Chimeric BVDV	Mari-viral synth	Combined open	(/ Hepatitis	Compliance of	6 Composite	Hepatitis C	cDNA sequ	0 40	Hepatitis C	7 Sense strand	4 Compiled HCV	.0 Hepatitis C viru	HCV polyprotein	Hebatitis C	20	4 Hepatitis C	HCV-1 NS3/4	Hepatitis C virus	ONA ENCOALD	Nucleotide	iah 69	Hepatitis C	1 Hepatitis C vi	S CDNA encoding	Hepatitis C vi	Mepatitis C	Mucleotium 10 Hepatitis	7 Hepatitis C	NS4A-NS3 com	ALIGNMENTS							n protease coding sequence #3.		muce10;		
B II		н	_	-	٦.	Α,		٦.	. 4	S.	0 (> c		0	٠.	٦ ,	,		_	m	~ c	D C	•			毋・	.	, 0	0	0	٦ .	٠ م	4	σ.	.	1 💠	4	20 AAX 24 AAD	4			94 BP.			.	3 fusion	viral repli	cer; mar		
Length Di		594	594	588	594	594	20 m	1000	12734	612	5300	2360	7310	7310	7310	8316	9185	9185	9185	9400	9401	9401	9401	9502	549	2058	2028	8145	9185	9185	1033	9646	96	23	16622	22	2061	1998 9365				DNA; 5			st entry	NS4A-NS	tease;	er con		
% Query Match	:		8	9	٠.	ς,				5	۲.	87.4	٠.	7	~ 1	~ r	87.4	~	\sim	~	7	٠,	۲.	۲.	7	ς,	٠,	۲,	7	7	٠, ۷		ė	٠			٠,	86.7 86.5				standard			(111	C virus	NS3 pro	ָ בּ	C virus	
Score	1	1032	1015	866	866	985	5/5	959	30	22	55	901.5	5	0	5	35	50	0.1	5	5	22	35	50	18	668	668	50 G	668	96	898.5	ž «	96	96	96	89.00 0.00 0.00	, α	ထ	894.5 892.5			_	330	73330;	•	-DEC-2000	itis	epatitis;	110	Hepatitis (Synthetic.	
Result No.		-	7	٣	₹ 1	'n	0 r	· 60	6	2:	1:	13	14	15	16	1 E	13	20	21	22	53	4 °C	26	27	58	5 6	30	35	33	34	3,50	37	38	6 °	4.4	4 5	43	44 45		RESULT 1	AAA73330	Æ,	AC AAA7	•	5	DE Hepat	KW Hep			

CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC 360

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                    /product= "NS4A-NS3 fusion protein #3"
                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                 Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                           Claim 26; Fig 13; 66pp; English.
                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                   Weinheimer S,
                                                                                                                                                                               99US-0115271
                                                                                                                                             2000WO-US00345
                    /*tag=
                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
P-PSDB; AAB15221.
                                                                       MO200040707-A1
                                                                                                                                                                                                                                                   Wittekind M,
                                                                                                                                             36-JAN-2000;
                                                                                                                                                                               08-JAN-1999;
                                                                                                          13-JUL-2000
```

Sequence 594 BP; 103 A; 186 C; 156 G; 149 T; 0 other; Alignment Scores

OTTAINER OF STEE	5	. 00				
Pred. No.:			2.42e-85	Length:	594	
Score:			1032.00	Matches:	197	
Percent Similarity:	imila	arity:	100.00%	Conservative:	0	
Best Local Similarity:	1 Str	nilarity:	100.008	Mismatches:	0	
Query Match:	ch:		100.00%	Indels:	0	
DB:			21	Gaps:	0	
13-09-965	-594	-14 (1-197	US-09-965-594-14 (1-197) x AAA73330 (1-594)	(1-594)		
οy	-	MetLysLys	LysGlySerValV	alileValGlyArgile	MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla	20
qa	1	ATGAAAAA	VAAAGGATCCGTTG	TTATCGTCGGCCGTATC	ATGAAAAAAAAGGATCCGTTGTTATCGTCGCCGTATCAACTGTGCCGGTGACACTGTCCCGTGACACCTGTCCGGTGACACTGTCGCGTGATCGTTGTTATCGTTCGT	09
٥y	21	TyrAlagl	GlnThrArgGlyG	luGluGlyCysGlnGlu	TyralaginginThrargglygluglugiyCysglngluThrSerGlnThrGlyArgasp	40
qq	61		SCAGACTCGAGGTG	aggaggettgccaaga?	TACGCTCAGCAGACTCGAGGTGAGGGTGGCCAGAACCTCCCAGACGGTCGTGAC	120
Qy	41		ValGluGlyGluV	alGlnIleValSerTh	LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla	09
q	121		GCTGAAGGTGAAG	TTCAGATCGTTTCCACC	AAAACGAGGTTGAAGGTGAAGTTCAGATCGTTTCCACGCGGTGCTCAGACCTTCCTGGCT	180
Qy	61		sAsnGlyValCysT	rpThrValTyrHisGl	ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla	80
qq	181		CAACGGTGTTTGCT	GGACCGTTTACCACGG	ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCGTACCATCGCT	240
ογ	81		sGlyProVallleG	InMetTyrThrAsnVal	SerProLysGlyProVallleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp	100
q	241		AGGTCCGGTTATCC	AGATGTACACCAACGT		300
Oy	101	ProAlaPro	oglnglySerArgs	erLeuThrProCysTh1	ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 	120

LeuvalThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140 CIGGITACCCGICACGTGACGITAICCCGGTTCGTCGTGGTGGTGACTCCCGTGGTTCC 420 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C LeuleuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys CIGCIGTCCCCCCCGGTCCCACCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGTGTGTCT protease; viral replication; chronic liver disease; liver cancer; mutant; mutein; ds. AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197 591 Hepatitis C virus NS4A.NS3 fusion protease coding sequence #4. /*tag= a /product= "NS4A-NS3 fusion protein #4" Goldfarb V; Weinheimer S, Zhang Y, Location/Qualifiers (BRIM) BRISTOL-MYERS SQUIBB CO ВÞ 06-JAN-2000; 2000WO-US00345. 990S-0115271 AAA73331 standard; DNA; 594 (first entry) 1..594 /*tag= WPI; 2000-465976/40. P-PSDB; AAB15222. Hepatitis C virus. Hepatitis; NS3 liver failure; WO200040707-A1 Wittekind M, 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Synthetic AAA73331; 141 121 361 181 421 RESULT a ò QQ ð g ò 셤 ò

Claim 26; Fig 14; 66pp; English.

The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.

m

```
121 LeuvalThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                               P-PSDB; AAB15220
      WO200040707-A1
                                                                                                                                                                                                                                                                                             Sequence 588
                                                                                      Wittekind M,
                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                            Alignment Scores:
                       13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
à
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                     140
                                                                                                                                                                                                                                                                                                                    420
                                                                                                                                                                                                                                                                                                                                                                              240
                                                                                                                                                                                                                                     100
                                                                                                                                                                                                                                                                     120
                                                                                                                                                                                                                                                                             160
                                                                                                                                                                                                                                                                                                                                                     480
                                                                                                                                                       120
                                                                                                                                                                                      180
                                                                                                                                                                                                                                                     ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaI,ys 180
                                                                                                                                                                      9
                                                                                                                                                                                                      80
                                                                                                                                       9
                                                                                                       20
                                                                                                                      9
                                                                                                                                                                                                                                    CTGGTTACCCGTCACGCTGACGTTATCCCGGTTCGTCGTCGTGGTGACTCCCGTGGTTCC
                                                                                                                                                                                                                                                                                                                                              LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                     ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                             ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                      LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                     LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                               1 ATGAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGT
                                                                                                                                       TyralaGlnGlnThrArqGlyGluGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; muteln; ds.
                                                                                                                                                                                                                                                                                                                                                                                                    181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product= "NS4A-NS3 fusion protein #2"
       other;
                              594
1194
0
0
0
0
                                      Matches:
Conservative:
Mismatches:
Indels:
      0
      Ë
      147
                              Length:
                                                                                       US-09-965-594-14 (1-197) x AAA73331 (1-594)
      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
      155
                                                                                                                                                                                                                                                                                                                                                                                                                                                     BP
       ပွဲ
                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA73329 standard; DNA; 588
                              8.61e-84
1015.00
98.48%
98.48%
      A; 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /*tag=
      BP; 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus.
                                                      Best Local Similarity:
Query Match:
DB:
                                               Percent Similarity:
      594
                       Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic
        Sequence
                                                                                                                                       21
                                                                                                                                                       61
                                                                                                                                                                      41
                                                                                                                                                                                      121
                                                                                                                                                                                                      61
                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                      241
                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                     141
                                                                                                                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                                                                                                      161
                                                                                                                                                                                                                                                                                                                                                                                     481
                                                                                                                                                                                                                                                                                                                                                                                                                     541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA73329
                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA73329
AAA73329
AAA73329
AAA73329
AAA73329
AAA73329
AAA73229
DI 19·D
XX
XX
HEPA
XX
HEPA
XX
HEPA
XX
XX
HEPA
XX
HEPA
XX
XX
XX
HEPA
HIVE
                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                    οχ
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                Š
                                                                                                       ò
                                                                                                                     셤
                                                                                                                                       ò
                                                                                                                                                      셤
                                                                                                                                                                   õ
                                                                                                                                                                                   g
                                                                                                                                                                                                     ά
                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                  δ
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both resential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2
                                                                                                                                                                                                                                                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophillic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerProLysGlyProValIleGlnMetTyrThrasnValaspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CCGGCTCCGCAGGCTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAAAACCAGGTTGAAGGTGAAGTTCAGATTCGTTTTCCACCGCTGCTCAGACCTTCCTGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 ATGAAAAAAAAGGATCGTTGTTATCGTCGCCGTATAGTACTGAACGGT-----GCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            588
193
1
1
2
1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   other
                                                                                                                                                             Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP; 103 A; 180 C; 156 G; 149 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-14 (1-197) x AAA73329 (1-588)
                                                                                                                                                          Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 26; Fig 12; 66pp; English.
                                                                                                         ဗ
                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3.03e-82
998.00
98.48%
97.97%
06-JAN-2000; 2000WO-US00345
                                                      990S-0115271
                                                                                                                                                             Weinheimer S,
                                                                                                                                                                                                                WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
```

AAA73335

q õ a

õ

ò

```
US-09-965-594-14 (1-197) x AAA73335 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BP.
              8.79e-82
993.00
97.46%
97.46%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAA73332 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1..594
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
Synthetic.
                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
     Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAA73332;
                                                                                           Query Match:
                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA73332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               òγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify protease and protease:inhibitors complexes. The protein produced from this sequence contains the alpha-helixO wild-type sequence.
                                                                                                                            ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                   AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein #8"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "NS4A-NS3 fusion
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ပ္ပ
                                                                                                                                                                                                                                                                                                         BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                       AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /*tag= a
/product=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-465976/40.
P-PSDB; AAB15226.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                 181
                                                                                                                                                                                                                                    535
                                                                                                                                                                                                                                                                                                                                          AAA73335;
                                                                                         415
                                                                                                                            191
                  355
                                                       141
```

Key

```
140
                                                                                                                                                                                                                                                                                                                                                                                 420
                                                                                                                                                                                                                                                                                                                                                                                                       160
                                                                                                                                                                                                                                                                                                                                                                                                                           480
                                                                                                                                                                                                                                                                                                                                                                                                                                                ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       540
                                                                                                                                                                                                        180
                                                                                                                                                                                                                                                                        100
                                                                                                                                                                                                                                                                                            300
                                                                                                                                                                                                                              80
                                                                                                                                          40
                                                                                                                                                                                    9
                                                                                                                    9
                                                                                                                                                                                                                             LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                      CTGCTGTCCCCGGGTCCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTGGTGCTGCTGTGTGCTGTGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       481 CCGCCTGCTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA
                                                                                                                     1 ATGAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                                                                                     41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                               LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                              MetLysLysLysGlySerValVallleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                          21 TyralaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /*tag= a
/product= "NS4A-NS3 fusion protein #5"
Conservative:
Mismatches:
            Matches:
  Length:
                                            Indels:
                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W0200040707-A1
```

S

us-09-965-594-14.rng

```
19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic
                                        421
                                                                                161
                                                                                                                      481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                         NAME OF THE PROPERTY OF THE PR
                                     g
                                                                                ò
                                                                                                                    a
                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80
                                                                                                                                                                                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21 TyralaglnglnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyargasp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ThrCys1leAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCTCCATCAACGGTGTTCTGTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   594
191
0
6
0
                                                                                                                                                              Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-14 (1-197) x AAA73332 (1-594)
                                                                                                                                                              Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                  Claim 26; Fig 15; 66pp; English.
                                                                                                                    (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                              Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.72e-81
985.00
96.95%
96.95%
                                        2000WO-US00345
                                                                                99US-0115271
                                                                                                                                                                                                    WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                           P-PSDB; AAB15223
                                                                                                                                                              Wittekind M,
                                        06-JAN-2000;
                                                                                08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                              ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence is the coding sequence for a mutated version of a
                                  141 LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                         Hepatitis C'virus NS4A-NS3 fusion protease coding sequence #6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /*tag= a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              594
188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ဥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5.88e-80
973.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-JAN-1999;
```

ø

us-09-965-594-14.rng

```
08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
  ò
                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                         οž
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                       ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaVs 180
                                                                                                                                                                                                                                                                                      540
                                                                                                                                                240
                                                                                                                                                            SerProLysGlyProValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                   140
                                                                                                                                                                                                                               420
                                                                                    9
                                                                                                                                  80
                                                              9
                                                                                                                                                                                                                                                                                      CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA
                                               LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                          CIGGITACCGICACGCIGACGITATCCCGGTTCGTCGTCGTCGTGGTGACTCCCGTGGTTGC
                                                                            TyralaglnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                       LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                  ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                   AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A·NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           /*tag= a
/product= "NS4A-NS3 fusion protein #7"
 2100
 Conservative:
       Mismatches:
Indels:
                     Gaps:
                                   (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
1..594
                                   US-09-965-594-14 (1-197) x AAA73333
                                                                                                                                                                                                                                                                                                                                            ВР
                                                                                                                                                                                                                                                                                                                                            AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000 (first entry)
96.45%
95.43%
94.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                          /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
Synthetic.
        Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
Percent Similarity:
                                                              ٦
                                                                                                                                                                                                                               361
                                                                                                                                                                                                                                                                                                                541
                                                                                         61
                                                                                                       4.1
                                                                                                                     121
                                                                                                                                  19
                                                                                                                                                                           241
                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                             141
                                                                                                                                                                                                                                                                        191
                                                                                                                                                                                                                                                                                    481
                                                                                                                                                                                                                                                                                                                                                        AAA73334;
                                                                                                                                                                                                                                                          421
                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                             81
                Query Match:
                                                                                                                                                                                                                                                                                                                               RESULT 7
                                                                                                                                                                                                                                                                                                                                      AAA73334
                                                                                                                                                                                                                                                                                                                                            ò
                                                              a
                                                                             ò
                                                                                         셤
                                                                                                       ò
                                                                                                                    a
                                                                                                                                  õ
                                                                                                                                              a
                                                                                                                                                             ò
                                                                                                                                                                          g
                                                                                                                                                                                         ð
                                                                                                                                                                                                    a
                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                          đ
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                 셤
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCICCATCAACGGIGITCTGIGGACCGITTACCACGGIGCTGGIACCCGIACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LysasnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      594
187
20
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative:
Mismatches:
Indels:
                                                                             Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-14 (1-197) x AAA73334 (1-594)
                                                                             Zhang Y,
                                      8
                                                                                                                                                                                                                                                                       Claim 26; Fig 17; 66pp; English.
                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.82e-79
963.00
95.94%
94.92%
93.31%
                                                                             ŝ
99US-0115271
                                                                             Weinheimer
                                                                                                                 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                 P-PSDB; AAB15225
                                                                           Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
```

161

```
161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SerProLysGlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                  1 MetLysLysLysGlySerValVallleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                            21 TyralaGlnGlnThrArqGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArqAsp
                                                                                                                                                                                                                                                                                41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                            61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla
                                                                                                              1 ATGAAAAAAAAGGTTCCGTTGTTATCGTCGGCCGTATAGTACTGAACGGT----GCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pestivirus; Npro; protease; NS3; screening; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
     US-09-965-594-14 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               04 - DEC - 2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABA95615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hong Z,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    aba9561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THE STATE OF THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                  ò
                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3 mytants fullon proteins which can be used to identify inhibitors of this type, as
                               ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                ВЪ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protease:inhibitor complexes
                                                                                                                                                                                                                                                                                AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1.1e~78
959.00
95.94%
95.43%
92.93%
                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1..588
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P-PSDB; AAB15212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JB-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUL-2000
                                                       481
                                                                                                                                                                                                                                                                                                                                  AAA73328;
```

Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)

Pred. No.:

294

234

80

9

40 54

œ

```
protease gene, useful for in vivo screening of compounds which inhibit HCV infection
pestivirus genome where the Npro protease gene is replaced with NS3
```

Example 2; Columns 17-28; 20pp; English.

Anti-viral synthetic prototoxophore associated DNA sequence.

(first entry)

28-MAR-2003

ABX15706;

chimeric Hepatitis C virus (HCV) pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimmeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention. present invention relates to a nucleic acid construct encoding a

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

```
ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                       LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CITICGCCCCGGCCCATITCCTACCTAAAGGCTCCTCGGGGGGGGTCCGCTGTTGTGCCCC 892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                                                                                            GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                         22 AlaGluGlnThrArgGlyGluGluGlyGysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                       42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                    Length:
Matches:
                                                                                                                                                                                                                             (1-12734)
                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                              US-09-965-594-14 (1-197) x ABA95615
                                                                                                                                                                  1.94e-74
930.50
94.87%
92.82%
90.16%
                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                              122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                142
                                                                                                                                                                                                                                                                                                                                                                                                                                                              773
                                                                                                                                                                                                    Query Match:
                                                                                                                                                                              Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                8
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                              ò
ò
                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                             셤
```

```
This invention relates to a novel synthetic viral prototoxophore comprising a toxin molety operatively incorporated into a substrate domain specific for a viral enzyme. This prototoxophore may be bound and modified by the viral enzyme thus converting it to a toxophore. Also disclosed in the invention is a method for enhancing the anti-viral effect of an antiviral agent, this method comprises contacting a cell, infected with a virus or is susceptible to infection, with a prototoxophore. The invention further comprises an assay to identify anti-viral agents, comprising contacting an infected cell with a candidate agent and comparing the ability of the agent to inhibit the growth or infectivity of the virus in the cell. The prototoxophores of the invention may be useful for reducing or inhibiting viral infectivity, by contacting a cell (e.g. lymphocyte, nerve cell, connective tissue cell, muscle cell or hepatocyte) which is infected with a virus or is susceptible to infection with a virus, with a virus of its susceptible to infection with a virus, with a virus of the prototoxophore is also useful for ameliorating the severity of a viral infection in a subject, where the virus is selected from human immunodeficiency virus (Hist), herpes simplex virus (HSV), rhinovirus and hepatitis virus, by administering an effective amount of the prototoxophore to the subject. The prototoxophores of the invention are antiviral prototoxophore associated DNA sequence represents an antiviral prototoxophore associated DNA sequence. The prototoxophore is a necental DNA sequence, this sequence is a networthy and received by administering and prefer sequence represents an antiviral prototoxophore associated DNA sequence. The prototoxophore is a recombinant with NSJ/NS4 fusion protein in example 1 of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          has
                                                                                                                                      Hepatitis C; ds; viral prototoxophore; anti-viral; tumour;
virus; infection; antitumour; toxophore; human immunodeficiency virus;
HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Novel synthetic viral prototoxophore for treating viral infections, toxin molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into toxophore
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the invention although it is clearly not a protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ħ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Shepard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 1; Page 62; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                        26-APR-2002; 2002WO-US13223.
                                                                                                                                                                                                                                                                                                                                                                                                                 27-APR-2001; 2001US-286893P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Neuteboom STC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                              (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                            WO200287500-A2
                                                                                                                                                                                                                               Unidentified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Cathers BE,
                                                                                                                                                                                                                                                                                                                          07-NOV-2002
```

532

4,

592

61

652

81

612 181 3 8 8 1

Length:
Matches:
Conservative:
Mismatches:
Indels:

Percent Similarity: Best Local Similarity:

Score:

Query Match:

Gaps: (1-612)

US-09-965-594-14 (1-197) x ABX15706

BP.

standard; DNA; 612

ABX15706

RESULT 10 ABX15706 ID ABX1

953

Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other

07e-75 921.50 94.36% 92.82% 89.29%

Alignment Scores:

σ

```
927 TIGCIGGCGCCCATCACGCCTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAAIC 986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   34 ThrSerGlnThrGlyArgAspLySAsnGlnValGluGlyGluValGln1leValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 867 CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     987 ACCAGCCTAACTGGCCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTGTCAACT
                                                                                            Claim 3; Figure 26-1, 26-2, 26-3, 26-4, 26-5, 26-6; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                           Sequence 5300 BP; 1047 A; 1606 C; 1515 G; 1130 T; 2 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 LysLysGlySerValVallleValGlyArglleAsn-----
                                                                             and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                  (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to correct PI field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-14 (1-197) x AAN92097 (1-5300)
                            Kuo G;
                                                                                                                                                                                                                                                                                                                                                                                                                     2.93e-72
                                                                                                                                                                                                                                                                                                                                                                                                                            901.50
90.20%
85.78%
87.35%
                                                                     Purified hepatitis C virus
                            Choo QL,
          (CHIR ) CHIRON CORP
                                            WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity:
                                                     P-PSDB; AAP92041
                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similari
                            Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                    Pred No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                 ProLysGlyProValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 101
                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                             498
                                           121
                                                                                                                                                                                                                                            LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                          open reading frames of the hepatitis C virus (HCV) cDNA in
4i llb, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c,
                                                                                                      81
         61
                                                                                                              CCCAAGGGTCCTGTCATCCAGATGTATCCAATGTAGACCAAGACCTTGTGGGCTGGCCC
                                                                                                                                                                                                           ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                        AlabroGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                    AsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThr
                                                                                                     CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                  22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                              ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                               559 GTGGACTTTATCCCTGTGGAGACCTAGAGACAACCATGAGGTCC 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
3..5300
                                                                                                                                                                                                                                                                                                                                                                 AAN92097 standard; DNA; 5300 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87US-0122714.
87US-0139886.
88US-0161072.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                   (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                            8f, 33f, 33g and 39c.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-MAY-1988;
26-OCT-1988;
14-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                          02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                     lones 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                  AAN92097;
                                                                                    139
                                                                    42
                                                                                                     62
                                                                                                                                                       59
                                                                                                                                                                        03
                                                                                                                                                                                         319
                                                                                                                                                                                                           122
                                                                                                                                                                                                                                           142
                                                                                                                                                                                                                                                                                                               182
                                                                                                                                                                                                                                                                                                                                                                                                                           Combined
                                                                                                                                                                                                                                                                                                                                                RESULT 11
                                                                                                                                                                                                                                                                                                                                                          AAN92097
                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                          g
                                                                   ò
                                                                                  g
                                                                                                     õ
                                                                                                                  ద
                                                                                                                                  à
                                                                                                                                                      q
                                                                                                                                                                                        셤
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                               a
                 q
                                   ò
                                                   a
                                                                                                                                                                        ò
```

us-09-965-594-14.rng

Sun Aug

```
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-MAR-2003
02-MAR-1990
                                                                                                                                                                                    867
                                                                                                                                                                                                                                      34
                                                                                                                                                                                                                                                                                         1047
                                                                                                                                                                                                                                                                                                                           1107
                                                                                                                                                                                                                                                                                                                                                            1167
                                                                                                                                                                                                                                                                                                                                                                                                                                1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                  154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAN92103;
                                                                                                                                                                                                                                                                                                                                            94
                                                                                                                                                                                                                                                                                                                                                                                                                134
                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAN92103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BXANKAXB
                                                                                                                                                                                  q
                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                         òγ
                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
  8888888
                                                                                                                                                                   ö
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                        ò
                                                                               CGGGGTGATAGCAGGGGCAGCCTGCTGTCGCCCGGCCCATTTCCTACTTGAAAGGCTCC 1346
                                                                                                                                                          rcggggggrccgcrgrrgrcccccgcgggcaccccgrggcararrraggccgcgcgrg 1406
                    GCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTA 1166
                                                                                                        153
                                                                                                                                          SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                      CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArg 133
  6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The sequence shows the composite cDNA sequence derived from the aligned hepatitis C virus (HCV) cDNA's in clones 14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f, 33f, 33g and 39c. The cDNA
                                                                                                                                                                             CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               prodn. of polynucleotide probes, diagnosis, prevention and treatment
AlaGlyThrArgThrIlcAlaSerProLySGlyProValIlcGlnMetTyrThrAsnVal

        ArgGlyAspSerArgGlySerLeuLeuSerProArgPro1leSerTyrLeuLysGlySer

                                     AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                  Hepatitis C virus; composite cDNA; probe; vaccine.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Fig. 26; 174pp; English.
                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
3..5360
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               for
for
                                                                                                                                                                                                                                                                                                                          Hepatitis C virus composite probe
                                                                                                                                                                                                                                                             ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polypeptide(s) and antibodies of infection.
                                                                                                                                                                                                                                                                AAN90327 standard; cDNA; 5360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87US-0122714.
87US-0139886.
88US-0161072.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88US-0263584
                                                                                                                                                                                                                                                                                                                                                                                                                                                  88GB-0027024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (CHIR ) CHIRON CORPORATION.
                                                                                                                                                                                                                                                                                                  (updated)
(first entry)
                                                                                                                                                                                                                       ATGAGGTCCCCG 1478
                                                                                                                                                                                                             MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Choo OL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 1989-215054/30.
                                                                                                                                                                                                                                                                                                                                                            Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            26-OCT-1988;
                                                                                                                                                                                                                                                                                                 25-MAR-2003
11-NOV-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Houghton M,
                                                                                                                                                                                                                                                                                                                                                                                                               GB2212511-A
                                                                                                                                                                                                                                                                                                                                                                                                                                26-JUL-1989
                   1107
                                                    1167
                                                                                       1227
                                                                                                        134
                                                                                                                        1287
                                                                                                                                                         1347
                                                                                                                                                                                            1407
                                                                                                                                                                                                                             1467
74
                                    76
                                                                      114
                                                                                                                                         154
                                                                                                                                                                           174
                                                                                                                                                                                                            194
                                                                                                                                                                                                                                                                               AAN90327
                                                                                                                                                                                                                                                       AAN90327
                                                                                                                                                                                                                                                                 셤
                                   ŏ
                                                    8
                                                                      ò
                                                                                      g
                                                                                                        ç
                                                                                                                        셤
                                                                                                                                         õ
                                                                                                                                                         셤
                                                                                                                                                                            å
                                                                                                                                                                                           셤
                                                                                                                                                                                                             ç
                                                                                                                                                                                                                              셤
à
```

```
SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                       :::::||||
CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The CDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to raise antibodies for immnosasay or treatment, or to produce vaccines. See also AAP90158, AAN90303-26, and AAN90328-36.
                                                                                                                                                                                                                                                                                                                                                                                                  54 AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrH1sGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                        Sequence 5360 BP; 1060 A; 1622 C; 1532 G; 1145 T; 1 other;
                                                                                                                                                                                                                                                                                                                                   3 LysLysGlySerValVallleValGlyArgIleAsn------
                                                                                                                                                                         5360
175
9
                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                      (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                 (1-5360)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BP.
                                                                                                                                                                                                                                                                                                 US-09-965-594-14 (1-197) x AAN90327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAN92103 standard; DNA; 6905
                                                                                                                                                                       2.97e-72
901.50
90.20%
85.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MetArgSerPro 197
                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
```

1619

153

133

```
Combined open reading frames of the hepatitis C virus (HCV) cDNAs from clones K9-1 through 15e.
                                                                     1500 TGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCGG
                                                                                                                                                                                 154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
 74 AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal
                                                         AspLysAspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                             114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVallleProValArgArg
                                                                                                                                                                 134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 3; Figure 47-1 - 47-8; 139pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers 3..7310
                                                                                                                                                                                                                                                                                                                                                                                                                   AAN92106 standard; DNA; 7310 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ΰ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          88EP-0310922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kuo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Purified hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                   1740 ATGAGGTCCCCG 1751
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /*tag=
                                                                                                                                                                                                                                                                                                                                    MetArqSerPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1989-159274/22.
P-PSDB; AAP92050.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18-NOV-1987;
30-DEC-1987;
26-FEB-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-NOV-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25-MAR-2003
02-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-MAY-1988
26-OCT-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          14-NOV-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31-MAY-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EP318216-A.
                                                                                                                                                                                                                                                                                                                                    194
                                                                                                                                                                                                                                                                                                                                                                                                                                               AAN92106;
                                                        8
                                                                                                                                                                                                                                                                                                                                                                                           RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                         AAN92106
                           a
                                                                                  a
                                                                                                            ò
                                                                                                                                      g
                                                                                                                                                                 à
                                                                                                                                                                                          g
                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1140 CGCAGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1200 TIGCIGGCGCCCAICACGGCGTACGCCCAGCAGAGAGGGCCTCCTAGGGTGCAIAAIC 1259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus (HCV) cDNa. It can be used to make oligomeric DNa hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodies useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 6905 BP; 1421 A; 2082 C; 1946 G; 1456 T; 0 other;
                            Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 LysLysGlySerValVallleValGlyArglleAsn.....
                                                                                                                                                                                                                                                                                                                                                                                                                  Purified hepatitis C virus - and associated nucleic acids and polypeptide(s)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 3; Figure 32-1 - 32-7; 139pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  JS-09-965-594-14 (1-197) x AAN92103 (1-6905)
                                                                                Location/Qualiflers
3..6905
/*tag= a
                                                                                                                                                                                                                                                                                                                                                 Kuo G;
                                                                                                                                                                                                                    87US-0122714.
87US-0139886.
88US-0161072.
88US-0191263.
88US-0263584.
88US-0271450.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4.05e-72
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                            88EP-0310922
clones 12f through 15e
                                                                                                                                                                                                                                                                                                                                                Choo QL,
                                                                                                                                                                                                                                                                                                                      (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                           WPI; 1989-159274/22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                        P-PSDB; AAP92047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                Houghton M,
                                                                                                                                                                                           18-NOV-1988;
                                                                                                                                                                 31-MAY-1989
                                                                                                                                                                                                                    18-NOV-1987
                                                                                                                                                                                                                                   30-DEC-1987
26-FEB-1988
                                                                                                                                                                                                                                                                06-MAY-1988
                                                                                                                                                                                                                                                                             26-OCT-1988
                                                                                                                                                                                                                                                                                          14-NOV-1988
                                                                                                                                      EP318216-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score:
```

셤 à g ò a

à

g

ò

(updated)

```
2024
                                                                                                                                                                                                                                                                                                                                                                                                                                          It is a double-stranded nucleotide sequence of the open reading frame (ORF) (tag a) extending through clones K9-1 to 15c of hepatitis C virus (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to detect the presence of HCV nucleic acids in samples. The polypeptide(s) it encodes could be used as immunoassay reagents and vaccines and to generate antibodles useful in diagnosis and passive immunotherapy for HCV infection/non-A, non-B hepatitis.
(Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                        CGCAGGGGCCGGGAGATACTGCTCGGGCCAGCCGATGGAATGGTCTCCAAGGGGTGGAGG 1724
                                                                                                                                                                                                                                                                                                                                                                                                  2145 ressesserecererrersececsessessessessessessessesrerrasses
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AspLysAspLeuvalGlyTrpProAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               93
                                                                                                                                                                                                                                                                                                                                                                                     33
                                                                                                                                                                                                                                                                                                                                                                                     ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaalaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            74 AlaGlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GACCAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTGACACCCTGCACT
                                                                                                                                                   Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 4 other;
                                                                                                                                                                                                                                                                                                                             3 LysLysGlySerValValIleValGlyArgIleAsn----
                                                                                                                                                                                             Length:
Matches:
Conservative:
                                                                                                                                                                                                                                        Mismatches:
Indels:
                                                                                                                                                                                                                                                                                              (1-7310)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ВР
                                                                                                                                                                                                                                                                                                US-09-965-594-14 (1-197) x AAN92106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAN90336 standard; DNA; 7310
                                                                                                                                                                                      4.35e-72
901.50
90.208
85.788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2265 ATGAGGTCCCCG 2276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MetArqSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (updated)
                                                                                                                                                                                                                                         Similarity:
                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25-MAR-2003
                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1785
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1845
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                194
                                                                                                                                                                                                                                                                                                                                                                                     12
                                                                                                                                                                                                                                                                                                                                                                                                                                             34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAN90336;
                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAN90336
                                                                                                                                                                                                            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
    88888888888888
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
1844
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1725 TTGCTGGCGCCCATCACGCGTACGCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATC 1784
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by aligning clones k9-1 through 15e in 5'-3' direction. The cDNA encodes antigens which react with antibodies in patients with non-A non-B hepatitis (NANBH). The cDNA can be used to design probes, or to synthesise polypeptides, which are used to diagnose HCV-induced NANBH, to raise antibodies for immunoassay or treatment, or to produce vaccines. See also AAP9028B, and AAN90303-35.

(N.B. This record was resubmitted to correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR field.)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             74 AlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus gene - used for prodn. of polynucleotide probes, polypeptide(s) and antibodies for diagnosis, prevention and treatment of infection.
                                                                         ds.
                                                                      Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 LysLysGlySerValVallleValGlyArgIleAsn---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-14 (1-197) x AAN90336 (1-7310)
                                         Composite hepatitis C virus (HCV) cDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; fig 47; 235pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.35e-72
901.50
90.20%
85.78%
87.35%
                                                                                                                                                                                                                              87US-0139886
88US-0161072
                                                                                                                                                                                      88GB-0027024
                                                                                                                                                                                                                 87US-0122714
                                                                                                                                                                                                                                                            88US-0263584
                                                                                                                                                                                                                                                                                                                     Kuo
                                                                                                                                                                                                                                                                                       (CHIR ) CHIRON CORPORATION.
                (first entry)
                                                                                                                                                                                                                                                                                                                   Houghton M, Choo QL,
                                                                                                                                                                                                                                                                                                                                              WPI; 1989-215054/30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                             P-PSDB; AAP90288
                                                                                                  Pan troglodytes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
19-JUL-2001
01-NOV-1989
                                                                                                                                                                                    .8-NOV-1988;
                                                                                                                                                                                                                                                            26-0CT-1988;
                                                                                                                               GB2212511-A
                                                                                                                                                                                                                              30-DEC-1987,
26-FEB-1988,
                                                                                                                                                           26-JUL-1989
                                                                                                                                                                                                                 18-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred.
 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
```

```
1905 GCCGGAACGAGGACCATCGCGTCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTA 1964
   1965
a
   ò
```

Search completed: August 30, 2003, 19:47:57 Job time : 192.939 secs

```
Compugen Ltd.
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compug
```

OM protein - nucleic search, using frame_plus_p2n model

August 30, 2003, 19:20:43; Search time 1910.31 Seconds (without alignments) 2506.388 Million cell updates/sec Run on:

1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197 BLOSUM62 Xgapop 10.0, Xgapext (Ygapop 10.0, Ygapext (Fgapop 6.0, Fgapext (Delop 6.0, Delext US-09-965-594-14 Title: Perfect score: Scoring table: Sequence:

0.5

22781392 seqs, 12152238056 residues Searched:

45562784 of hits satisfying chosen parameters: Total number

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Command line parameters:
-MODEL-frame+_p1.mcdel-DEV=xlp
-MODEL-frame+_p1.mcdel-DEV=xlp
-MODEL-frame+_p1.mcdel-DEV=xlp
-Cgn2_1/USPTC_spool/US09965594/runat_29082003_151919_28322/app_query.fasta_1.2872
-DG-Cgn2_1/USPTC_spool/US09965594/runat_29082003_151919_28322/app_query.fasta_1.2872
-DG-ALIGN-200 -TRE_SCORE-pct -THR_MATRIX-DIOSUM62 -TRANS-buman40 -Cd1 -LIST-45
-UGCRGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL
-UGCRG-USD-GCG -THR_SCORE-PC -THR_MAX-100 -ALIGN-15 -MODE-LOCAL
-UGCRG-USD-GCG - HASPIZE-SOO -MINIEN-200000000
-UGCRG-USD-GCG - LI_12630_drunat_2908203_151919_28322 -NCPU-6 -ICPU-3
-NO_MATR -LIREGOUERY - NEG_SCORES-0 - WAIT -DSPBLCCK-100 -LONGLOG
-DEV_IMBOUT-120 -WARN_IMBOUT-30 -THREADS-1 -XGAPEXT-0 -S -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

EST:* Database :

em_gss_rod:* em_gss_phg:* em_gss_vrl:* gb_gssl:* em_gss_hum:* em_gss_inv:* em_gss_vrt:* em_gss_fun:* em_gss_mam:* em_gss_mus:* em_gss_pln:* em_gss_pro:* em_esthum:
em_estin:
em_estov:
em_estov:
em_estpl:
em_htc:
gb_estl:* em_estfun:* em_estom: * gb_est2:* gb_htc:* gb_est4:* gb_est5:* em_estba:* gb_est3:* 14: 15: 17: 19:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	escription	4699 601888	2487 AGENCO	1625 BJ	4121 BJ0241	6176 BJ0161	3316 601865	6101 AGENCO	0545	1437	3244	2274	7233	80	1887	824	8485	4657	9745	2790	7051	284	424	707	240	280	200	125	10	CB286751 CMD45 C08	614	081	142	586	553	687	980	318	560	841	551	418	7	18059	76374	
SUMMARIES	Ü	9469	248	16	412	617	331	610	3054	143	324	3227	723	908	88	824	98	465	760	5279	3705	284	5424	2201	9 4 5 0 6 0 6 0 6 0 6	0861	2002	100	1000	CB286751	5614	BU54081	BX45142	CD10586	BQ67229	BUS5687	BG42086	BQ67318	BI75560	BG17841	BG38551	BF79418	BQ9282	BF18059	CB76374	
	DB	34 10	1 66	<u> 5</u>	13 1	54 1	51 1	33 1	11	79 1	11 1	16 1	71	6 70	3.		4	P S		01	32	05	66	16	00.	10	4.6	יי	4 4	82 14	15	46 1	66 1	81 1	29 1	47 1	58 1	79 1	08 1	71 1	98 1	63 1	0	40	22 1	
	ery tch Lei	0.3	.0	9.6	. بو	9.	.5	.5	.5	₹.	٣.	۳.	۳.	7	~ •	7	7.0	7	⊣ -:	٦.	٦.	٦.		- -:	د د	٠. د د	· ·		•	0.6	. 0.	0.	0.	0.	0.	0.	0.	0.	0.	6.	6.	σ.	.9	6.	σ.	
	ore	106	'n	66		9		8	σ			6			σ,	4	2. 2.	•	۷.	m i	m	m,	33.5	ສ໌		20.0	m 0	2 6	٠,	92.5		ď	ď	ď	'n	ď	ď	'n	ď.	92	92	92	92	σ	91.5	
	Suli			3		5	9 0		æ	on	10	c 11	12	13	4 1	٠,	0 1 1 1	٠,	0 18	19		c 21	~ (C 73	~ (71 (40	3 0	4 (30	m	32	33	34	35	36	c 37	38	c 39	40	c 41	42	4	C 44	4	

ALIGNMENTS

984 bp mRNA linear EST 21-NOV-2000 601888252F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4122276 5', BF304699 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 984) BF304699.1 GI:11251586 EST. Homo sapiens (human) Homo sapiens BF304699/c LOCUS DEFINITION ORGANISM ACCESSION VERSION KEYWORDS SOURCE REFERENCE RESULT 1

```
source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                          REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 СD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     뎝
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                        /organism="Homo sapiens"
/mol_type="mRNA"
/Mol_type="mRNA"
/do_type="mRNA"
/do_type="thadGe:412276"
/tissue_type="thadGomyosarcoma"
/lab.host="hull0B (phage-resistant)"
/lab.host="hull0B (phage-resistant)"
/lab.host="hull0B (phage-resistant)"
/lab.host="hull0B (phage-resistant)"
/lab.host="hull0B (phage-resistant)"
/lab.host="hull0B (phage-resistant)"
/note="Organ: muscle: Vector: porB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5, adaptor: GGGGGGGG(G). Size-selected >500bp
for average insert size 1: Rkb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-CDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1199 bp mRNA linear EST 16-AUG-2002
AGENCOURT_8417538 Lupski_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:6192708 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119 LeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100 TrpProAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     553 GGGGGCTCTTGTGGGGAAGACCTCGATGGTGTCCAAGCTCGCGCTGCTGTACTGGAAGT 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           493 CGGCACGCTCCGGTCAGGTGCAGC--------TTCCAGCGCCCGGGG 455
                                                                                   Email: cgapbs-rémail.nih.gov
Tissue Procurement: ATC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Incytte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1065 row: g column: 13
High quality sequence stop: 646.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------ACCAGGCACGCAGCGAACATACATGCAAGGAGACGTGGT---TCCCGC
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        984
333
224
522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches:
Conservative:
Mismatches:
Indels:
                                                                Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-14 (1-197) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BQ892487
BQ892487.1 GI:22284501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         106.00
45.24%
39.29%
10.27%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 CysThrArgGly 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TGCGCCCGAGGA 443
                                                                                                                                                                                                                                                                                             1. .984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                298
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BQ892487
                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
//Essue_type="sympathetic trunk"
//esv_stage="adult, 16 yr"
//dev_stage="adult, 16 yr"
//dev_stage="adult, 16 yr"
//deb.host="bnil08"
//deb.host="bnil08"
//deb.host="bnil08"
//deb.host="bnil08"
//deb.host="bnil08"
//desp_fice="bnil08"
//des
                                                                                                                                                                                                                I Unpublished
Contact: Robert Strausberg, Ph.D.
Email: capabs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium
CDNA Education: MGC clone distribution information
Clone distribution: MGC clone distribution information
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.lln.gov
Plate: LLAM13595 row: c column: 13
High quality sequence start: 57
High quality sequence stop: 394.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----CysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         143 -----SerProArgProlleSerTyrLeuLys----------GlySerSerGl 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88 GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySerArg 107
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. (bases 1 to 1199) NIH-WGC http://mgc.nci.nih.gov/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68 TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValIle 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          484 TGGGATCCATTTTTAATAAAGGGTGCTCTGTTAATCATGGCGCCCACGGGCCCGCTGATA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          574 ---TGTGCTGCCTGCACAGCACCCCCATGACCGATGTTGGGGCTTATGTGGAAACGGCGAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 119 LeuTyr-LeuValThr-----ArgHisAlaAspValIleProValArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches:
Conservative:
Mismatches;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   211 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Homo sapiens'/mol_type="mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-14 (1-197) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /db_xref-"taxon:9606"
/clone="IMAGE:6192708"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            echnologies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              108 SerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /sex-"male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10.4
103.50
37.42%
26.45%
10.03%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
```

Db 373ACTGTGGAGGAGGAGGAGGTGTAGG 335	Qy 112 CysThrCysGlySerSerAspLeuTyrLeuValThrArg 124	Oy 125HisalaaspValileProValArgArgArgGlyAspSer 137	0y 138 ArgGlySerLeuLeuSerProArgProlleserTyrLeuLysGlySerSer 154 1 1 1 1 1 1 1 1 1	Qy 155 GlydlyProLeuLeuCysProAlaGlyHisAla 165 	RESULT 4 BJ024121 COCUS BJ024121 COC	BJ04411 MF0155A CUNA OLYZIAS IACIPES CUNA MRNA Sequence. BJ024121 BJ024121 GI:17377389	SOUNCE OTYZIAS Intipes (Japanese medaka) ORGANISM OTYZIAS Jatipes Characoa; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Teleosteti; Euteleosted; Actinopterygii; Neopterygii; Teleosteti; Euteleosteli; Actinopterygii; Neopterygii; Teleosteli; Neopterygii	REFERENCE 1 (bases 1 to 643) William School of the American School o	COMMENT CONTACT: Ideasus Sinn-1 Center For Genetic Resource Information National Institute of Genetics Illl Yata, Mishims, Shizuoka 411-8540, Japan	TELT OF 2559-81-6855 FRAX: 81-559-81-6855 Email: tshini@genes.nig.ac.jp. FEATURES Location/Qualifiers		/ db_xrei="taxon:8090" / clone="MF01SSA143D12" / sex="mixture of female and male" / tissue_type="whole embryo" / dev_stage="segmentation stage 20 - 25" / clone_lib="M01SSA cDNA" BASE COUNT 171 a 148 g 176 t	Alignment Scores: 12.7 Length: 643 Pred. No.: 99.00 Matches: 42 Score: 99.00 Matches: 9 Best Local Similarity: 27.81% Mismatches: 50	9.59% Indels: 12 Gaps: -197) x BJ024121 (1-643)	Qy 41 LysAsnGlnValGluClyGluValGlnIleValSerThralaAlaGlnThrPheLeuAla 60	242 AAAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT	l Qy 61 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
Db 751 AAAACCCCACCCACCGATGGGTTAAGCTCCCTTTACAAGCCACCGGGGG 810	Oy 155 yGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysTh 175 1 1 1 1 1 1 1 1 1	Qy 175 rArgGlyValAlaLysAlaValAspPhelleProValGluSer 189 :::	RESULT 3 BJ001625/c LOCUS BJ001625 G15 bp mRNA linear EST 05-DEC-2001 DEFINITION BJ001635 MEDISCA CHAR OFFICER CHARACTER ACTORS 6.	MRNA sequence. N BJ001625.1 GI:17364516	<pre>REIMOUS EST. SOUNCE Oryzias latipes (Japanese medaka) ORGANISM Oryzias latipes Dukaryotas Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi;</pre>	Acuthoprerygii; Neoprerygii; Isleostei; Luceleostei; Neotergostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Beloniformes; Adrianichthyidae; Oryziinae; Oryzias. REFERENCE I (bases I to 615) AUTHORS Kohara,Y., Shin-I.T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.	TITLE MEGRAR EST Project in Takeda S lab JORNAL Unpublished COMMENT Contact: Tadasu Shin-i Conter For Genetic Resource Information National Institute of Genetic or	Anatomat Mishima, Shizuoka 411-8540, Japan 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-606	Source 1.013 /organism="Oryzias latipes" /mol_type="mrna" /strain="Hd-rR" /db. woof="too.ooon"	/ub_xrel= caxolineovo /clone="MFOlSSA025C02" /sex="mixture of female and male" /fissue_trype="whole embryo" /dov etane="semmentation etane 20 - 25"	,	Alignment Scores: Pred. No.: Score: Score: Score: Store: Best Local Similarity: 27.81% Ouery Match: 12 Length: 615 Matches: 42 Percent 33.77% Conservative: 9 Best Local Similarity: 27.81% Mismatches: 50 Query Match: 12 Gaps: 7	US-09-965-594-14 (1-197) x BJ001625 (1-615) Qy 41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 60 Lillil 1 ::: Db 511 AAAAAGAGGAAGAAGAAGAACAAGATCCAACGAACATGTTCTACGGGGG 452	Qy 61 ThrCyslleAsnGlyValCySTrpThrValTyrHisGlyAhaGlyThrArgThr11eAla 80 1 1 1 1 1 1 1 1 1	Qy 81 SerProLysGlyProVallleGlnMetTyrThrAsnValAspLys 95	409 GCTCTGCGGCGCGAGGACTCCTGGGCCAGTTGTG	Qy 96 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro 111

```
EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 6
BF203316/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                 셤
                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          ç
                               ò
                                                           음
                                                                                      ò
                                                                                                                                                ò
                                                                                                                                                                          a
                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                            BJ016176 BF01SSA CDNA Oryzias latipes CDNA clone MF01SSA025C02 3',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Oryzias latipes
Oryzias latipes
Dukaryota: Metazoa, Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii: Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
                                                                                                                                                                                                                                                                                          530 AGAGGCCGGACAGGCAGTCGGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCA 589
                                                                                                                                                                1 (bases 1 to 754)
Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
-------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGCCA 343
                                                           96 AspLeuValGiyTrpProAlaProGlnGlySerArgSerLeuThrPro------ 111
                                                                                                                    -----ACTCGTGGAGGACGAAGAGCGTCACCCGGGAGCTTAGG 418
                                                                                                                                                                                                        470 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGC 529
                                                                                                                                                                                                                                                               ArgGlySerLeuLeuSerProArg------ProlleSerTyrLeuLysGlySerSer 154
                                                                                                                                                112 ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg-------
                               81 SerProLys ------GlyProValIleGlnMetTyrThrAsnValAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 754
42
9
50
50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unpublished
Contact: Tadasu Shin-i
Contact: Tadasu Shin-i
Contact: Tadasu Shin-i
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Fax: 81-559-81-6856
Fax: 81-559-81-6855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /clone="MF015SA025C02"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20
/clone_lib="MF01SSA cDNA"
181 c 181 g 198 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                        155 GlyGlyProLeuLeuCysProAlaGlyHisAla 165
                                                                                                                                                                                                                                                                                                                                         /organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /db_xref="taxon:8090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1. .754
                                                                                                                                                                                                                                                                                                                                                                                                                                       mRNA sequence.
BJ016176
BJ016176.1 GI:17376695
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            33.778
27.818
9.598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15.6
99.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ರ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                      125
                                                                                                                                                                                                                                                                                                                                                  590
                                                                                                                    380
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                               RESULT 5
BJ016176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score:
                                                                                                                                                                                                                                                                                    đ
                                                           a
                                                                                      ò
                                                                                                             g
                                                                                                                                                ð
                                                                                                                                                                   q
                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                        ö
```

```
BF203316 961 bp mRNA linear EST 06-NOV-2000 601865914F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4098578 5',
                                                                                                                                                                                                                                                                                                                                                                                                                    344 GCTCTGCGGCGGCGAGGACCTCCTGGGCCAGTTGTG-------379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     96 AspLeuValGlyTrpProAlaProGlnGlySerArgSerLeuThrPro------ 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           380 ------ACTCGTGAGGACGAGGACGTCACCCCGGAGCTGTAGG 418
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     112 ------CysThrCysGlySerSerAspLeuTyrLeuValThrArg------ 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                419 CIGCAGGGAIGCGGAIGTGGCTCTGCT-----TTGGTTCCTCGTCCTCGGATCA 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         470 TCTTCCTCACCTGACCTTCCACATCCAGGTGTCGCCAGGGCTGGTCTGACGGGTGATGGG 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           530 AGAGGCCGGACAGGCAGTCGGGGTGAATCTCTGCAGGACGTCTTCACGCCGGATCA 589
                                                                                                                                                                                                                                                                                                   302 -----------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     138 ArgGlySerLeuLeuSerProArg------ProIleSerTyrLeuLysGlySerSer 154
                                                                                                                                               242 AAAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT 301
                                                                         41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAla 60
                                                                                                                                                                                                                      61 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla 80
                                                                                                                                                                                                                                                                                                                                                                         81 SerProLys------GlyProVallleGlnMetTyrThrAsnValAspLys 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota, Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 951)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Contact: Robert Strausberg, Ph.D.
Email: Gapbs-remail.nih.gov
Tissue Procurement: ArCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incytte Genomics, Inc.
Clone distribution: MGC chone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM965 row: I column: 03
High quality sequence stop: 637.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /clone="IMAGE:4098578"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/lone_lib="NH_MGC_17"
/note="Organ: muscle; Vector: pOTB7; Site_1: EcoR1;
Site_2: Xhoi; cDNA made by oligo-dT prining.
Directionally cloned into EcoRI/Xhoi sites using the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     155 GlyGlyProLeuLeuCysProAlaGlyHisAla 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1. .961
/organism-"Homo sapiens"
US-09-965-594-14 (1-197) x BJ016176 (1-754)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type-"mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BF203316.1 GI:11096902
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     mRNA sequence.
BF203316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unpublished
```

S

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               65
                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                      ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
JOURNAL
MEDLINE
                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
AK080545
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AGENCOURT_8752655 NIH_MGC_130 PmRNA linear EST 20-AUG-2002
5', mRNA sequence.
BQ926101
BQ926101.1 GI:22341132
EST.
              for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." 297 c 300 g 134 t
                                                                                                                                                                                                                                                                                                                                                                                                                                            160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             160 ysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaL 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1403)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unpublished
Contact: Robert Strausberg, Ph.D.
Email: gapbs-remail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
c)NA Library Preparation: ResGen, Invitrogen Corp
cDNA Library Arrayed by: The L.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13798 row: j column: 07
High quality sequence stop: 101.
                                                                                                                                                                                                                                                                                                         101 ProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer-AspLeuTy
                                                                                                                                                                                                                                                                                                                          120 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe
                                                                                                                                                                                                                                                                                                                                                                                                                                            140 rLeuLeuSer-ProArgProIleSerTyrLeuLysGlySerScrGlyGlyProLeuCeuC
                                                                                                                                                                                                                                                                                                                                                                                                          591 GACTCTGAGTGCAGGAGCGGACAGAGATGCGGATCGGAAGCGCGGGGATGCTGATGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             477 GTCCTGCAGTGCTGCTTGTAGGTGCTGTCCGTCAGCTTTTTGTGATGGTGGTGACGCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          531 CATGCTGTCCTCCTCGTCCGCGGGAGTAGGCTAGGGTCTGGAGGAGGAGGG----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism-"Mus musculus"
/mol_type-"mRNA"
/dolaref-taxon:10090"
/clone="IMAGE:6335718"
/lab_host-"DH10B (phage-resistant)"
                                                                                                                                                    961
32
11
32
7
following 5' adaptor: GGCACGAG(G)
                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                       US-09-965-594-14 (1-197) x BF203316 (1-961)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus (house mouse)
                                                                                                                                                  23.8
98.50
52.44%
39.02%
                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               180 ys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               417 AA 416
                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                B0926101/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                    SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
```

a

ò

셤

g ò a

õ

```
AKO80545

AKO80545

Mus musculus 7 days neonate cerebellum cDNA, RIKEN full:length enriched library, clone:A730082L10 product:weakly similar to zinc finger protein (fragment) [Mus musculus], full insert sequence.

AKO80545

AKO80545. I GI:26348600

HTC; CAP trapper.

Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ||||:::
| 1321 IGTCGGGGCGCGTTGGCGCATACCCCGGGTCGGATCGAGGTCAGCGCCGCTTGTATACA 1262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1249 GGGGTA---TGGTTATCACGGGCTGGGGCAGGTACT------TCCCCTAAAAGCG 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1378 GGGTGTGTTCANCGGTCAGGACAGGTCGCC---GCACACTCGACGCGTCGGCCAGAGACT 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -----CAG 1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                84 lyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProG 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      137 ----SerArgGlySer-----LeuLeuSerProArgProlleSerTyrLeuLysG 152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       45 GluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCyslleAsn 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11 GlyArglleAsnLeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGly 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  104 InGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    124 rgHisAlaAspValIleProValArgArgGlyAsp------
/clone_lib="NIH_MGC_130"
/note="Organ: otcoysts; Vector: pCMV-SPORT6.1.ccdb;
/note="Organ: otcoysts; Vector: pCMV-SPORT6.1.ccdb;
Priner: Oligo dT. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is
NIH_MGC Library.
521 c 237 g 345 t 3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys--G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31 CysGlnGluThr ---- SerGlnThrGlyArgAspLysAsnGln -----
                                                                                                                                                                                                                                                     1403
49
12
62
43
                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279233
                                                                                                                                                                                                                                                                                                                                                             Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-14 (1-197) x BQ926101 (1-1403)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           152 lySerSerGlyGly 156
                                                                                                                                                                                                                                                        38.5
98.50
36.75%
29.52%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GAGAAGGCGGCGGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1261 GAGGGGAAA----
                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                         297
```

```
/translation="bsclpaaspgrslltprgdgfflkeklsaaravgpgpsvargvs
Yrkordadggegorrvcaarrsbsclltbrrrrrchvppvgphyyglhyglsggrripppage
AqaacrapqyvPhpsprhggyvPpggaapptlbalaarqvpvgpyrgrrgprraprhs
pkpvptalgfsrggggpapplllapangrsvglll"
                                                                                                                                                                                                                                                                                                                                                                                                                                                 /db_xref="FaNTOM_DB:A730082L10"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone="A730082L10"
/tissue_type="corebellum"
/clone_lib="RIKBN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in RIKEN contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86 ValileGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGly 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----CCTCCTCAAGGA 389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     126 AlaAspValileProValArg---ArgArgGlyAspSerArgGlySerLeuLeuSerPro 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            145 ArgProlle------SerTyrLeuLysGlySerSerGlyGlyProLeu 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----CAAGIGCCGCATCCGCCIGGG--------CGGCCACAIGGAACA 371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             106 SerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHis 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         159 LeuCysProAla---GlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGly 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46 GlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCyslleAsnGly 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66 ValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 390 GCGGCTGGGCTTCTCCCTGCA-------CTCGCAGCTCGCCAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note="unnamed protein product; putative weakly similar to zinc finger protein (fragment) musculus) (PIR I148722, evidence: FASTY, 50.7%ID, 57.6%length, match=601)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1141
46
11
52
50
                                                                                                                                               prepare mouse tissues.
Please visit our web site for further details.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             303 GGAGAAGCGCAGGCAGCTGGACGCGCGCGCGCAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches:
                                                                                                                                                                                                            URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /protein_id="BAC37940.1"
/db_xref="GI:26348601"
                                                                                                                                                                                                                                                                                                                                            /organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-14 (1-197) x AK080545 (1-1141)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note-"putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /note="putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         372 GTGGTG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  316 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33
98.00
35.85%
28.93%
9.50%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyA_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyA_site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
                                                                                                                                                                                                                                                                                   FEATURES
   COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
Arawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
Fleischman,W., Casterland,T., Gissi,C., King,B., Kochiwa,H.,
Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G., Tomita,M.,
Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
Carninci,P., de Bonaldo,M.F., Brownstefn,M.J., Bult,C.,
Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
Marchionni,L., Mashim,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
Toyo-oka,K., Wang,K.H., Weltz,C., Whittaker,C., Wilming,L.,
Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The FANTOM Consortium and the RIKEN Genome Exploration Research Analysas of the mouse transcriptome based on functional annotation of 60,770 full-length cobas

Nature 420, 563-573 (2002)

Nature 420, 563-573 (2002)

Rature 420, 563-573 (2002)

Recompleted as a farman and a farman annotation of 60,770 full-length cobas

Recompleted as a farman and a farman and a farman annotation and a farman 
                                                                                                                                                                                                                                                                            Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishi,I.Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watshiki,M., Yonaca,Y., Ishikawa,T., Ozawa,K., Tamana,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format Genome Res. 10 (11), 1757-1771 (2000)
          Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (16.4PR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-reségsc.riken.go.jp, ORL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9212,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Functional annotation of a full-length mouse CDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Muramatsu, M. and Hayashizaki, Y.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1217851
10349636
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11076861
                                                                                                                                                                                                    JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      JOURNAL
MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE
PUBMED
REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
PUBMED
                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
                                                                                                                                        TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
```

~

```
BFB63244 19-JAN-2001
963042C02.xl C. reinhardtii CC-1690, Stress condition I, normalized
, Lambda Zap II Chlamydomonas reinhardtii cDNA, mRNA sequence.
BFB63244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          118 -AspLeuTyrLeuValThrArg-----HisAlaAspValIleProVal-----ArgAr 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  229 TGACAGGCGCGGGGGGGGCGCCCATGCCGGGCCCGCGGCGGTCGTTCTTCGCCA 288
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            133 gArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySe 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          289 CCGCGGCCGTTGCCGGGGCACTATACG TCGCCGGGGCACGACGAGGAGAAGAACGGG 347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ------GlyHisA 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -------GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpPro 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    109 CCCTGTTCTGCCAGGTGGCCGCGGTGGACGCGCGCGCAGGGAAGCGGCTCGTGGTCG 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  348 crecegresecercecaraceaccessacrecesresecersesecercesecars 407
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 CAGCTGCCAACGCGGCCCCGCGAACAACTACAGGACGGTGC------TGCTGG 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             69 ThrvalTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVallle--- 87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chlamydomonas reinhardtii
Chlamydomonas reinhardtii
Bukaryota; Viridiplantes; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadaceae; Chlamydomonas.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 (bases i to 701)
Grossman, A., Davles, J., Federspiel, N., Harris, E., Hauser, C.,
Lefebvre, F., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D.
Analyses of the Chlamydomonas reinhardtil Genome: A Model,
Unicellular System for Analyzing Gene Function and Regulation in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               102 AlaProGlnGlySerArgSer------LeuThrProCysThrCysGlySerSer
see close 1J, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for bariley genomics. Bariley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)*
216 311 g 105 t
                                                                                                                                                                                                                                                            779
47
14
63
30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      153 rSerGly-GlyProLeuLeuCysProAla-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 165 laValGlyIlePheArgAlaAlaValCysThrArg 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GCGGAGGACGCGCGCGGGCTCTGCGT 442
                                                                                                                                                                                                                                                                                                                           Conservative:
                                                                                                                                                                                                                                                                                                                                                            Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (1-779)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-14 (1-197) x BF631437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BF863244.1 GI:12253388
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Contact: Charles Hauser
DCMB Box 91000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Durham, NC 27708-1000
Tel: 919 613 8159
                                                                                                                                                                                                                                                        22.7
97.50
39.87%
30.72%
9.45%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Duke University
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unpublished
                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
                                                                                                                                 ø
                                                                                                                              147
                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88
                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KEYWORDS
SOURCE
ORGANISM
                                                                                                                          BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                            ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BF863244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE
                                                                                                                                                                                                                                                               Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /clone_lib=_notial
/clone_lib=_notial
/clone_lib=_notial
/clone_lib=_notial
/clone_lib=_notial
/clone_lib=_notial
/clone_lib=_notial
/clone_lib=_notial
/clone_lib=_notial
/clone_notial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Wagnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
Triticeae; Hordeum.
1 (bases 1 to 779)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.H., Fenton, Y.D., Oates, R. and Maln, D.
Development of a genetically and physically anchored EST resource for barley genomics: Morex drought-stressed seedling shoot cDNA
                                                                                                                                                                                                                                                        HVSMED0015P05f Hordeum vulgare seedling shoot EST 122-OCT-2001 PVCDNA0002 (Dehydration stress) Hordeum vulgare subsp. vulgare cDNA BF631437
Fenton). Phagemids were plated and ploked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see
                                                                                         178 ValAlaLysAlaVal-----AspPheIleProValGluSerLeuGluThrThrMet 194
                                                                                                                                                     503 GACTIGAGGGIGGITGGCICIGAGCIGACACCAAGCICCICICIGGCAGGIACACII 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      On Dec 19, 2000 this sequence version replaced gi:11895595.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /organism-"Hordeum vulgare subsp. vulgare"
/mol_type-"mRNA"
/cultivar-"Morex"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Emall: rwing@clemson.edu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /db_xref="taxon:112509"
/clone="HysMab0015p05f"
/tissue_type="Seedling shoot"
/lab_host="1JC121"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Total hg bases = 455
Seq primer: AATTAACCCTCACTAAAGGG
High quality sequence stop: 588.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hordeum vulgare subsp. vulgare
                                                                                                                                                                                                                                                                                                                                                                                                                                               BF631437.2 GI:13092107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      source
                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                             BF631437
LOCUS
                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
```

TITLE

COMMENT

œ

```
Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                         176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
                                                                                                                                                                                                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
9
                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
 AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BF307233
                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                 BF182274 846 bp mRNA linear EST 31-0CT-2000 601804028F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:4035102 5', mRNA sequence.
BF182274.1 GI:11060416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          171 CACCACCATACCCTTGCTCTCACGTTGCTCACAAAATTATGCCCATACGGGCCACTA 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           107 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         127 AspValIle........ProValArgArgArgGlyAspSerArg-.... 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  87 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 231 ACAAAGTTACATACACGG-----AAGGACCAGGGGTTGGCCACCCCCTTGGAGCGG 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      345 GAGATCATTIGCATGTGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTATAAA 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           405 GAAGGGGACGGGAATTCCGTTTGCGGAAAAGTGAGCGCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   71 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys------GlyProVal 86
                                                   /organism="Chlamydomonas reinhardtii"
/mol_type="mrRNA"
/strain="CC-1690 wild type mt+ 21gr"
/db_xref="taxon:3055"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       151 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               165 CICCAAGGCAGCAATGGGAGCCTTTCGCGGGTGTGCGGGT 503
                                                                                                                                                                                                                                                                                                                                                                701
32
14
14
22
                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-14 (1-197) x BF863244 (1-701)
                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mus musculus (house mouse)
Mus musculus
             Email: chauser@duke.edu
                                                                                                                                                                                                                                                                                                                                                               27.6
96.00
40.718
28.328
Fax: 919 613 8177
                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                           173
                                                                                                                                                                                                                                                                                                                                                    Aliqnment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
BF182274/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCUS
                                                                                                                                                                                                                                                                                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VERSION
KEYWORDS
                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

```
/crganism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/strain="CS7BL/6J"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:405102"
/tissue_type="tumor, gross tissue"
/dev_stage="T months"
/dev_stage="T months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Mam5"
/clone_lib="NCI_CGAP_Ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                709 GGTICTTCAGTCCCAAGTGGGCTGGAGAAGTGAACACAGAGAAAAGAGGGCGGTCCCTCA 650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          104 nGlySerArg-----SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVa 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142 uSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeu---LeuCysPr 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----CATCCTTC 710
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           161 oAlaGlyHis---AlaValGlyIlePheArg-------AlaAlaValCysThrAr 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                73 GlyAlaGlyThrArg-ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAs 92
                                                                                                                                                                                                                                                         Email: cgapbs-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Robinson Robinso
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              122 IThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               609 -ACTAGACACAT--GTAATCCAGGAGGAACGCNTGGAGGAACAGAGGGACTCC---CT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    757 GGTTCTGCTACCAGAACAGCCTGGATGAAGAAAGGACCA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               846
13
36
31
91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                   Unpublished
Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-14 (1-197) x BF182274 (1-846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     241 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   176 gGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             438 TGGGATGACTGAG 426
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              218 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        35.1
96.00
47.20%
36.80%
9.30%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1. .846
```

```
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
                                                                                                                                                                                                                                                                                                                                                                    Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J., Vallet, J., Wise, T., Schrer, G.A., Pertea, G., Sultana, R., Quackenbush, J. and Keele, J.W.

Porchine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
                                                                                                                                                                   EST 09-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 CAAACCCCGGGCAGCTCTTCGGCCGCATCCCTATGCGCTTCTCGGTGCTGATGCAGATGC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121 GTTTCGACGGCCTGCTGGGCTTTCCCGGGGCTTCGTGGACCGGCGCTTCTGGTCATTGG 180
  --- CGGGCATCTCTCCAGGGCCGCTGGACTG 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    --TyrHis-Gl 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    73 yAlaGlyThrArgThrIleAlaSerProLysGlyProVallleGlnMetTyrThrAsnVa 93
                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 407)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/labot="blost="blost"
/labot="blost" blost"
/note="Vector: pcMV SPORT6; Site_1: Not1; Site_2: Sal1;
/ibrary made from pooled tissue from day 11, 13, 15, 20,
a ni3 0 embryos."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 43 GlnValGluGlyGluValGlnIleValSerThrAlaAlaGlnThrPheLeuAlaThrCys
                                         160 sProAlaGlyHisAlaValGly-----IlePheArgAlaAlaValCys 174
                                                               AW785806 407 bp mRNA linear 117260 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence. AW785806.1 GI:7842582
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        407
41
3
52
52
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tal: 402 762 48390
Fax: 402 762 4390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63 IleAsnGlyValCysTrpThrVal------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism-"Sus scrofa"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-14 (1-197) x AW785806 (1-407)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCAGTACGACG
Plate: 37 row: D column: 16
Seq primer: ATTTAGGGGGACATATAG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
767 CCTTTTGGTCCTGA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31.218
29.088
9.218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Contact: Smith TPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         407
                                                                                                                                                                                                                                                                    Sus scrofa (pig)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                          Sus scrofa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12226715
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                                                                                                                                  DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment
                                                                                                                       RESULT 13
AW785806
                                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                                                                                                                                     LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1. 901
/ Organism="Memor sapiens"
/ Mol_type="Memor"
/ Mol_type="Memor"
/ Mol_type="Memor"
/ Mol_type="Memor"
/ Abc. xef="taxon:9606"
/ Clone="IbAGE:4137145"
/ IbC. host="DH108 (phage-resistant)"
/ Clone_lib="NIH_MGC_17"
/ Anote-"Organ: muscle; vector: poTB7; Site_1: EcoR1;
/ Anote-ToTAM made by oligo-dT priming.
Directionally cloned into EcoR1/XhoI sites using the following 5 adaptor: agcAcGAGGGG(0: Size-selected >SOUDP for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of Carlifornia, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BF307233 901 bp mRNA linear EST 21-NOV-2000 601891502F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4137145 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            140 rLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCy 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100 pProAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      80 aSerProLysGlyProvalIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTr 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ...... 731
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        66 ValCysTrpThr-ValTyrHis------GlyAlaGlyThrArgThrIleAl 80
                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 901)

NIH-MGC http://mgc.ncl.nih.gov/.

NIH-MGC http://mgc.ncl.nih.gov/.

Onpublished

Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                         Email: cgapDs-remail.nih.gov
Tissue Procurement: ATCC
CON Library Preparation: Ling Hong/Rubin Laboratory
CONA Library Arrayed by: The L.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics. Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1044 row: c column: 02
High quality sequence start: 6
High quality sequence stop: 684.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    901
37
8
28
45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-14 (1-197) x BF307233 (1-901)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       716 GCCCGCCCAICCGGG------
                                                                                  BF307233.1 GI:11254342
                                                                                                                           Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42.5
95.50
38.46%
31.62%
9.25%
                                           mRNA sequence.
BF307233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      =
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
Pred. No.:
                                                                                                          EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  source
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
ORIGIN
                                                          ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

g

ò

Q

a õ g ò

ΟŅ

```
AU127824 AU2RP2 Homo sapiens cDNa clone NT2RP2002160 5', mRNA
                                                                                                                                                                                                                           84 GlyProValIleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaPro 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             104 GlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr----- 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bages 1 to 641)
Ota T., Nishikawa T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
                                                                                                                                                                                                                                                                              30 GlyCysGlnGluThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGln 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10 ValGlyArgIleAsnLeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlu
                                                                                                                                                                                                                                                                                                                                                         50 IlevalSerThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCys----
                                                                                                                                                                                                                                                                                                                                                                                                                             68 --- TrpThrValTyrHisGlyAlaGlyThrArg-----ThrlleAlaSerProLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           460 GGCCCATCGGTC-----TTCCCC-----
                                   931
46
15
52
40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              135 GlyAspSerArgGlySerLeuLeuSerProArgPro 146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       582 TCTGACCAGCGGGGCACCTTCCCAGCTGTCT 617
                                                                          Conservative:
Mismatches:
                                     Length:
Matches:
                                                                                                                 [ndels:
                                                                                                                                   Gaps:
                                                                                                                                                                    US-09-965-594-14 (1-197) x BQ878887 (1-931)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AU127824
AU127824.1 GI:10988178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Isogai,T.
HRI human cDNA project
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (human)
                                   49.4
95.00
40.13%
30.26%
9.21%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens
                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sequence.
                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  478
                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
COMMENT
                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AU127824
                                                          Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                             đ
                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     931 bp mRNA linear EST 16-AUG-2002 AGENCOURT_8119707 Lupski_dorsal_root_ganglion Homo sapiens cDNA clone IMAGE:6177774 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Context: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://mage.lln.gov
Plate: LLANISSE row: e column: 07
High quality sequence stor: 18
High quality sequence stor: 18
High quality sequence stor: 705.
I. 931
I. 931

rce //dlanism*Homo sapiens*
//dolaref**RNNA*
//dlanism*Homo sapiens*
//colone="HARGE: 6177774"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.-GACTAGTTCTAGATCGCGAGCGCCCCT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life
                                                                                                             125
                                                                                                                                                   300
                                                                                                                                                                                        ----AlaAspVallleProValArgArgArgGlyAspSerArgGlySerLeuLeuSerPr 144
                                                                                                                                                                                                                                                              164
                                                                                                                                                                                                                                                                                  Eukaryota: Metazoa: Chordata; Craniata: Vertebrata; Euteleostomi; Mammalla; Eutheria: Primates: Catarrhini; Hominidae; Homo. | (bases 1 to 931) | NIH-MGC http://mgc.nci.nih.gov/. | NIH-MGC http://mgc.nci.nih.gov/. | National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /clone_lib-"Luppk!_dorsal_root_ganglion"
/note="Vector: pcMV-SPORT6 (Life Technologies); Site_l:
Not!; Site_2: Sall; cDNA made by oligo-df priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCG-3' and
                                         93 lAspLysAspLeuValGlyTrpProAlaProGlnGlySerArg----SerLeuThrPr
                                                                                                             111 oCysThrCysGlySerSerAspLeuTyrLeuValThrArgHis-------
                                                                                                                                                   T.....--GACCGAGGCCCGCACGCGCTCGTGGCCCTCTGTACGCGC
                                                                                                                                                                                                                           301 GGCAGCTGACGCTGCAGCTGCACGCCGTGGAGATCAGCGCGGGTGCACTCGCGCGACC
                                                                                                                                                                                                                                                              oArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHi
                                                                          -----GGGCTGCCTGCGCTCACGGAGCCAGACTACTTAAGCTCACACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /tissue_type-"dorsal root ganglia"
/dev_stage-"adult, 36 yr"
/lab_host-"DH10B"
181 AGGACGGACTGAACCGGGTGCTGGGCCT---GGGCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    183 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  253 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BQ878887.1 GI:22270895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Technologies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /sex-"male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      308 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ø
                                                                                                                                                                                                                                                                                                                                         164 s 164
                                                                                                                                                                                                                                                                                                                                                                            403 C 403
                                                                                                                                                                                          126
                                                                              215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                 BQ878887
```

459

83

```
1. 641
/organism="Homo saplens"
/mol_type="mRNN"
/db_xref="taxon:9606"
/clone="W12RP2002160"
/cell_type="teratocarcinoma"
/cell_line="W12"
/clone_lib="W12RP2"
/clone_lib="W12RP2"
/clone="Vector: PME18SFL3; mRNA from NT2 neuronal precursor cells after 2 weeks retinoic acid (RA) induction"
03 a 256 c 183 g 96 t 3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         170 rgAlaAlaValCysThrArg------183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87 IleGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpProAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           137 SerArgGlySer----LeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 154
                                                                                                                                                                                                                                                                                                                                                                                                                            117 SerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgGlyAsp 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                    310 CACGAC-----CCACAAGCCCACCCG------CCCACAGCCACCGCGACTTCC 351
                                                                                                                                                                                                                                                                                                                                               184 -- PhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   532 CAACCGTGCCCATCGTCAACCTCAAGGACGAACTGCTGTTTCCA 575
                                                                                                                                                                            641
14
14
56
39
                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Search completed: August 31, 2003, 04:27:28
Job time : 1914.31 secs
                                                                                                                                                                                                                                                              US-09-965-594-14 (1-197) x AU127824 (1-641)
Location/Qualifiers
                                                                                                                                                                            34.3
94.50
39.10%
30.13%
9.16%
                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                 103
                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                              Pred. No.:
                                                                                                                                 BASE COUNT
 FEATURES
                                                                                                                                             ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                    ò
```

```
Hepatitis C virus
                                                                                                                                                                                                                   ; Search time 44.5227 Seconds
(without alignments)
700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1: /SIDSI/gogdata/geneseqp-embl/AA1980.DAT: *
2: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1981.DAT: *
3: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1981.DAT: *
4: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1981.DAT: *
5: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1981.DAT: *
6: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1986.DAT: *
7: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1986.DAT: *
8: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1986.DAT: *
8: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1987.DAT: *
8: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1987.DAT: *
10: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1987.DAT: *
11: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1987.DAT: *
12: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1997.DAT: *
13: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1997.DAT: *
14: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1997.DAT: *
15: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1997.DAT: *
16: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1997.DAT: *
17: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1999.DAT: *
18: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1999.DAT: *
20: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1999.DAT: *
21: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1999.DAT: *
22: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1999.DAT: *
23: /SIDSI/gogdata/geneseq/geneseqp-embl/AA1999.DAT: *
24: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2000.DAT: *
25: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2001.DAT: *
26: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2002.DAT: *
27: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2002.DAT: *
28: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2003.DAT: *
24: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2003.DAT: *
25: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2003.DAT: *
26: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2003.DAT: *
27: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2003.DAT: *
28: /SIDSI/gogdata/geneseq/geneseqp-embl/AA2003.DAT: *
29: /SIDSI/gogdata/geneseq/ge
                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-16
1032
1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Description
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15222
AAB15221
AAB15223
AAB15224
AAB15220
AAB15225
AAB15225
AAB15226
AAB15226
AAB15226
                                                                                                                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                               August 30, 2003, 17:42:58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     %
Query
Match Length D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
                                                                                                                                                                                                                                                                                                                                                                                                                           Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ٠.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Database
                                                                                                                                                                                                                                   Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Result
```

HCV NS4A-NS3 COMP1 HCV NS4A-NS3 COMP1	HCV NS4A-NS3 COMP1	NS4A-NS	HCV NS4A-NS3 COMp1	HCV NS4A-NS3 compl	HCV NS4A-NS3 COMP1	HCV NS4A-NS3 COMPI	HCV NS4A-NS3 COMPI	HCV NS4A-NS3 COMP1	NS4A-NS3			HCV NS4A-NS3 COMP1		NS4A-NS3	NS4A-NS3	NS4A-NS3	NS4A-NS3	HCV NS4A-NS3 COMP1		NS4A-NS3	NS3 prote	atiti	Amino acid sequenc	Hepatitis C virus	HCV genomic amino	phCV150-encoded se	pHCV176-encoded se	ъ	Protein sequence o	Peptide encoded by		encode	encoded	
AAY 2494 AAY 2494	AAY2494	AAY178	AAY 2494	AAY2494	AAY2494	AAY 2494	AAY1788	AAY17		AAY24944														AAU9928	AAR4012		AAR7922	Ī	AAP9015	AAP9016	-	AAP92050	-	
20																																		
665 665	99	216	665	665	999	671	216	216	216	665	671	215	216	216	216	216	215	216	216	213	631	3011	3011	3012	3011	687	1648	1766	1786	2261	2301	2436	2436	2772
88.0 87.9	٠.					- 1	87.4	87.3	87.3	87.3	87.3	87.2	87.0	87.0	87.0	86.9	9.98	9.98	9.98	86.1	86.1	86.1	86.1	86.1	85.8	85.7	85.7	85.7	85.7	85.7	85.7	85.7	85.7	85.7
908.5	907.5	904.5	904.5	904.5	903.5	903.5	901.5	900.5	900.5	900.5	900.5	006	897.5	897.5	897.5	896.5	894	893.5	893.5	883	883	888.5	888.5	888.5	885.5	884.5	884.5	884.5	884.5	884.5	884.5	884.5	884.5	884.5
110	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

4	
z	
a	
É	
Z	
5	
4	
3	
ď	

```
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                             Goldfarb V;
                                                                                   Hepatitis C virus NS4A-NS3 fusion protease #4.
                                                                                                                                                                                                                                                                            Wittekind M, Weinheimer S, Zhang Y,
                     AAB15222 standard; protein; 197 AA.
                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                  99US-0115271.
                                                             19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                WPI; 2000-465976/40.
N-PSDB; AAA73331.
                                                                                                                                     Hepatitis C virus.
Synthetic.
                                                                                                                                                                   WO200040707-A1.
                                                                                                                                                                                                                                  08-JAN-1999;
                                                                                                                                                                                         13-JUL-2000
                                        AAB15222
RESULT 1
```

```
ö
                                                                                      The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                             TCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIPRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                    09
                                                                                                                                                                                                                                                                                                                                                                                                             09
substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                       DB 21; Length 197;
                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb V:
                                                                                                                                                                                                                                                                                                         100.0%; Score 1032; DB 21; 100.0%; Pred. No. 6e-101; Viennitches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15221 standard; protein; 197 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                              Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                  197; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                         197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUL-2000
                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT AAB1522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
g
                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
```

```
ö
                                                                                                            The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NG4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins valued to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 TCINGVCMTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGMPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 TCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protease; viral replication; chronic liver disease; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                    Score 1015; DB 21; Length 197; Pred. No. 3.8e-99;
                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                                                                                    0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBE CO
                                                                                  Claim 23; Fig 13; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                   98.4%;
98.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000MO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 98.5
Matches 194; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                     197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAA73332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15223;
                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

us-09-965-594-16.rag

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NG4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NG4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                               0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 TCINGVCWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                            9
            Modlfied hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                           1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                  Length 197;
                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                  Score 1002; DB 21;
Pred. No. 9e-98;
0; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15224 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                     Claim 23; Fig 15; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                   97.1%;
98.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 98.5
Matches 194; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                       197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15224;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local S
                                                                                                                                                                                                                                                                                variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NG4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NG4A well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              120
                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                               9
                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix \bf 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis \bf C
                                                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                  TCINGVCWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                              Length 197;
                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb V;
                                                                                                                                                                                                                                                                                                                              95.9%; Score 990; DB 21; 97.0%; Pred. No. 1.7e-96; Live 2; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15220 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                            Claim 23; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 97.0'
Matches 191; Conservative
                                                                                                                                                                                                                                                                                                    197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
N-PSDB; AAA73333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15220;
                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                             /ariant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 5
õ
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                      유
                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                             The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HeV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HVV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                             TCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                             LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                1 MKKKGSVVIVGRIVLNG--AYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 58
                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                               ;
                                                                                                                                                                                                                                                                                        Length 195;
                                                                                                                                                                                                                                                                                                             4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                       Score 981; DB 21;
Pred. No. 1.5e-95;
1; Mismatches 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15225 standard; protein; 197 AA.
                                                                                          Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                       95.1%;
96.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                             190; Conservative
WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                 195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
         N-PSDB; AAA73329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                      59
                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                   119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15225;
                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                           variant
                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
 õ
                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HOV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 MKKKGSVVIVGRINLSGDTAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protease; viral replication; chronic liver disease; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      95.0%; Score 980; DB 21; Length 197; 96.4%; Pred. No. 1.9e-95; ive 2; Mismatches 5; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease #8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15226 standard; protein; 197 AA.
                                                                                                                                                                                                                                  Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity 96.4
Matches 190; Conservative
2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
Synthetic.
                           N-PSDB; AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               variant
```

Ŋ

```
W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                                                                                                                                                                 complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY24943;
                                                                                                                                                                                                                                                                    Seguence
                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY24943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
   õ
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both assential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease.inhibitor complexes. This sequence contains the alpha-helix() wild-type sequence.
                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                           120
                                                                                                                                                                                                                                                                                                                                                                                               120
                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                                      9
                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                             61 TCINGVCWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDFAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                    LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK
                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                        DB 21; Length 197;
                                                                                                                                                                                                                                                                                                             Indels
 Goldfarb V;
                                                                                                                                                                                                                                                                                      94.6%; Score 976; DB 21; 95.9%; Pred. No. 5.1e-95; 1ve 0; Mismatches 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease #1.
Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ₹
                                                                                                         Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15212 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                  Best_Local Similarity 95.9
Matches 189; Conservative
                     WPI; 2000-465976/40.
N-PSDB; AAA73335.
                                                                                                                                                                                                                                                                   197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15212;
                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
AAB1521
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
```

```
61 TCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                               The present sequence is a fusion protein created using the Hepatitis C virus (HCV) N33 and N84A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of MS3 mutents and MS3-MS4 fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .;
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91.3%; Score 942; DB 21; Length 195; 93.9%; Pred. No. 2e-91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1; Mismatches
Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                           Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24943 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0094331.
Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches 185; Conservative
                                                          WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                          N-PSDB; AAA73328
Wittekind M,
```

9

```
(SCHE ) SCHERING CORP.
   (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                             665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                         Malcolm BA,
                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY24941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                          å
                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                        62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                             122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                      The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex domain, where the hydrophobic domain of native HCV NS3 berine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                    22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                              GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                   3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                          88.3%; Score 911.5; DB 20; Length 665; 86.7%; Pred. No. 1.9e-87; Live 16; Mismatches 7; Indels 3;
                                                                             New hepatitis C virus covalent complexes
                                                                                                   211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24947 standard; Protein; 665 AA.
                                   Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex SEQ ID NO:18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     102 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 86.7%
Matches 170; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                   Taremi SS,
           (SCHE ) SCHERING CORP.
                                                                                                    Claim 6; Page 90-92;
                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                      665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                07-SEP-1999
                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY24947;
                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                        S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX 24947
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                 ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 earlie protease domain. Where the hydrophobic domain of native HCV NS3 earlie protease domain. Where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NRR spectroscopy. They can also be used for detecting inhibitors by protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88.0%; Score 908.5; DB 20; Length
86.2%; Pred. No. 3.9e-87;
Live 17; Mismatches 7; Indels
   ż
                                                                                                                                  New hepatitis C virus covalent complexes
                                                                                                                                                                                                      Claim 6; Page 100-102; 211pp; English
Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24941 standard; Protein; 665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV NS4A-NS3 complex SEQ ID NO:12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 86.29
Matches 169; Conservative
Taremi SS,
                                                             WPI; 1999-385385/32
```

ï

us-09-965-594-16.rag

```
W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
 Malcolm BA,
                                                                                                                                                                                                                                             Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY17880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY17880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                      62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                   122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 earlie protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR specificacypy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                         5 GSVVIVGRINLSCD---TAYAQQTRCEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                      22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKITSLTGRDKNQVEGEVQVVSTATQSFLAT 81
                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                   3;
                                                                                                                                                                                                                                                                           87.9%; Score 907.5; DB 20; Length 665; 86.7%; Pred. No. 5e-87;
                                                                                                                                                                                                                                                                                                  8; Indels
                                                                                                                                                                                                                                                                                                  15; Mismatches
            Yao N;
                                                        New hepatitis C virus covalent complexes
                                                                              Claim 6; Page 85-87; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY24942 standard; Protein; 665 AA
            Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV NS4A-NS3 complex SEQ ID NO:13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                       Best Local Similarity 86.79
Matches 170; Conservative
            Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP
                               WPI; 1999-385385/32
                                                                                                                                                                                                                                                      665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            07-SEP-1999
            Malcolm, BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic
                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24942;
                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAX 24 942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                        NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 berine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminal of the HCV NS3 protease domain. The present soquence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by MR specificacy, They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3 The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             the non-covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ٠.
ش
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          87.9%; Score 907.5; DB
86.7%; Pred. No. 5e-87;
tive 15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Malcolm BA, Taremi SS, Weber PC, Yao N;
ż
                                                                                                                                New hepatitis C virus covalent complexes
                                                                                                                                                                                          Claim 6; Page 88-90; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY17880 standard; Protein; 216 AA.
Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity 86.7%
Matches 170; Conservative
Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (SCHE ) SCHERING CORP
                                                              WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               665 AA;
```

œ

```
Claim 6; Page 95-97; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY24946 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 86.2
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (SCHE ) SCHERING CORP.
WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                        665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24946;
                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                 NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 berine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VIRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                           5 GSVVIVGRINLSGD----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                      present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                .,
ن
                                                                                                                                                                                                                                                                                                                              87.6%; Score 904.5; DB 20; Length 216;
86.7%; Pred. No. 2.2e-87;
Live 16; Mismatches 7; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Yao N;
                                         New hepatitis C virus covalent complexes
                                                                      Claim 6; Page 76-77; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24945 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex SEQ ID NO:16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                Best Local Similarity 86.7
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (SCHE ) SCHERING CORP.
         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                     216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                        22
                                                                                                                                                                                                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24945;
                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY24945
```

a ç 셤 ò g ö a

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4A PST independent a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present complex: The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKITSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4a cofactor; NS4a-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87.6%; Score 904.5; DB 20; Length 665;
86.2%; Pred. No. 1e-86;
tive 16; Mismatches 8; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Yao N;
New hepatitis C virus covalent complexes
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                        New hepatitis C virus covalent complexes
                                                                          Claim 6; Page 97-99; 211pp; English.
```

665 AA; Sequence

3; Gaps Length 665; 8; Indels Query Match 87.6%; Score 904.5; DB 20; Best Local Similarity 86.2%; Pred. No. 1e-86; Matches 169; Conservative 16; Mismatches 8;

61 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT

62 CINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121

122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181

Q

à g

QQ ò

ò

182 VDFIPVESLETTMRSP 197

ò

Search completed: August 30, 2003, 19:12:23 Job time : 44.6227 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

mode1
S
using
search,
protein s
٠
OM protein .
ĕ

August 30, 2003, 18:01:52; Search time 9.75674 Seconds Run on:

(without alignments)
949.524 Million cell updates/sec

US-09-965-594-16 1032 Perfect score:

1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197 Sequence:

Gapop 10.0 , Gapext 0.5 **BLOSUM62** Scoring table:

127863 seqs, 47026705 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		ď			SUMMARIES	
Result No.	Score	Query Match	Query Match Length	DB	ID	Description
1	884.5	85	3011	-	POLG_HCV1	ιò
7	878.5	85.1	3011	-	POLG_HCVH	u P
æ	867.5	84.1	3010	٦	POLG_HCVTW	4
4	857.5	3	3010	-	POLG_HCVJT	29 h
2	853.5	ď	3010	٦,	POLG_HCVBK	ᅺ
9	853.5	82.7	3010	П	POLG_HCVJA	ב
7	677	5.	3033	~	POI.G_HCVJ6	h genome
60	675	65.4	3033	Н	POLG_HCVJ8	
6	98		321	٦	HHOA_ARATH	arabidops
10	84.5	8.2	485	Н	Y136_TREPA	~
11	81	•	452	Ч	AAMP_HUMAN	
12	78.5		437	Н	DEG1_ARATH	arabi
13	78	7.6	1705	٦	PTPO_MOUSE	mus n
14	76.5	7.4	264	-	CTRL_HUMAN	P40313 homo sapien
15	76.5	7.4	323	Н	VPRT_SMRVH	squir
16	76.5	7.4	333	٦	MOSA_RHIME	
17	76	7.4	911	Н	TB11_NEIMB	
18	75.5	7.3	209	Н	PAAD_PSEAE	psend
19	δ.	7.3	2663	_	CENE_HUMAN	bomod 4
20	75	7.3	388	-	ODPT_HUMAN	homo
21	75	7.3	455	٦	TMS5_MOUSE	mus m
22	75	7.3	594	٦	NIR_SPIOL	P05314 spinacia ol
23	74.5	7.2	106	-	TRFE_HORSE	
24		7.2	764	٦	ICCR_DROME	008180 drosophila
25	74	7.2	844	П	CN4A_RAT	
26	73.5	7.1	263	-	GRAK_MOUSE	5 mus m
27	73	7.1		7	IBP1_HUMAN	3
28	72.5	7.0		7	MLTD_ECOLI	P23931 escherichia
29	•	7.0		-	THYG_RAT	P06882 rattus norv
30	72	7.0		-	TRPD_PSEPU	P20575 pseudomonas
31	72	7.0		П	GAL1_STRCO	_
32	72	7.0		Т	POL_GALV	
33	72	7.0	-	Н	EGFR_MOUSE	•

MEROPS; S29.001; -.
IMEROPS; U39.001; -.
INCEPPCo; 1PR001410; DEAD.
InterPro; IPR002522; HCV_capsid.

EMBL; M62321; AAA45676.1; -. PIR; A39166; GNWVC3. PDB; 1A1V; 16-FEB-99. PDB; 1HEI; 25-NOV-98.

P11033 mus musculu Q90627 gallus gall P51124 homo saplen P31434 schizosacch P4883 oryctolagus Q9Rn5 frizobium 1 Q8pzr7 methanosarc Q01996 neisseria g P28863 oryctolagus P08359 felluk Q9hwc9 pseudomonas P13227 streptomyce
GRAD_MOUSE TRY1_CHICK GRAM_HUMAN CWG2_SCHPO ZP3_RABIT PRNE_HHILO HELS_WETWA TBP1_NEIGO MMO3_RABIT ENV_FLYGI RPOC_PSEAE GAL1_STRLI
нанананана
248 2548 2557 3557 326 730 730 1399 748
$\begin{array}{cccccccccccccccccccccccccccccccccccc$
71.5 71.5 71.5 71.5 71.5 71.7 71.7 70.5
W W W W W W W A A A A A A A A A A A A A

ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                01-406-1992 (Rel. 23, Created)
01-406-1992 (Rel. 23, Last sequence update)
01-506-1992 (Rel. 23, Last sequence update)
01-507-5003 (Rel. 42, Last annotation update)
6600me polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (N81); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/holicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate 1) (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -: CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-: CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [RNA](N).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN H AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
                                                                                                                                                                                                                                                                                                                                                                                                                                 [1]
SEQUENCE FROM N.A.
MEDLINE-91172826; PubMed-1848704;
Cloo Q.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,
Gallegos C., Colt D., Medina-Selby A., Barr P.J., Weiner A.J.,
Bradley D.W., Kuo G., Houghton M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -1 - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                  PRT; 3011 AA
                  STANDARD;
                                                                                                                                                                                                                                                                                    Hepacivirus.
NCBI_TaxID=11104;
              POLG_HCV1
P26664;
POLG_HCV1
```

```
a
                                                     d
                                                                             õ
                                                                                               g
                                                                                                                                                                                                                                                                                                                                         CELLULAR AMINOPEPTIDASE.
CARSID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN ( FOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                              RNA-DIRECTED RNA POLYMERASE (POTENTIAL) POTENTIAL.
                                                                                                                                                                                                                                                                                   Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 884.5; DB 1;
Pred. No. 4.3e-76;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GICNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                INTERPRO): IRRV0213; INV_ENIV.
INTERPRO): IRRV0213; INV_ENIV.
INTERPRO; IRR001818; HCV_NS3.
INTERPRO; IRR000145; HCV_NS4.
INTERPRO; IRR000145; HCV_NS4.
INTERPRO; IRR00190; HCV_NS4.
INTERPRO; IRR00190; HCV_NS5.
INTERPRO; IRR00156; HCV_RGRP.
INTERPRO; IRR00156; HCV_RGRP.
INTERPRO; IRR00156; HCV_RGRP.
INTERPRO; IRR001095; RNA_POI_DS_PS.
INTERPRO; IRR001094; RNA_POI_DS_PS.
INTERPRO; IRR001094; RNA_POI_PSVIT.
PFam; PF01543; HCV_CAPS; I.
PFam; PF01539; HCV_NS1; I.
PFam; PF01006; HCV_NS2; I.
PFam; PF01006; HCV_NS3; I.
PFam; PF01006; HCV_NS4; I.
PFam; PF01016; HCV_NS3; I.
PFam; PF01016; HCV_NS3; I.
PFam; PF010189; VITALIAGRP; I.
PFam; PF00271; helicase_C; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA; 327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                                                                                                                     3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ma
Local S.
172;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local S
Matches 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                               CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
```

```
1125 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 1184
                                                                                                                                     54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       115-2892 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 24). Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein E2 (GP68) (GP06) (NS1): Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.99-.); Protease/helicase NS3 (P70) (Hepacivitin)
(EC 3.4.21.98); Nonstructural protein NS4 (P4): Nonstructural protein NS4B (P77): Nonstructural protein NS5B (P66) (P77) (P77): Nonstructural Protein NS5B (P66) (P77): Nonstruct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
WEDLINE-98154321; PubMed-9493270;
Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Heperlitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
of unwinding:";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-97331322; PubMed-9187654; Yao W., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.; "Structure of the hepatitis C virus RNA helicase domain."; Nat. Struct. Biol. 4:465-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACTIVATION OF NGS.

-!- FUNCTION: NS54 SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-!- FUNCTION: NS54 IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICANTON.
-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Structure 6:89-100(1998).
-!- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-!- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBDINIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     υ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                     174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. MEDLINE-92052256;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus.
NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 {RNA}(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 2
POLG_HCVH
```

9; Mismatches

Conservative

Similarity

53

9: Gaps

Indels

m

```
POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 878.5; DB 1;
Pred. No. 1.6e-75;
                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                   N-LINKED (

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                85.1%;
83.3%;
                                                                                                                                                                                                                                                                                                                                                         480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity
                                                                                                                                                                                                                                                                                                 623
1232
1232
1232
1233
1234
1254
1258
1272
1291
1302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1323
1336
1352
1352
1352
1358
1373
1378
1382
1389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1410
1414
1419
1432
1438
1450
1456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11480
114880
115907
115907
115907
115908
115908
116098
11622
11622
11637
11637
                                      CARBOHYD
TUUN
CARBOHYD
CARBOHYD
TUUN
CARBOHYD
CARBOHY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HELIX
SEQUENCE
              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
STRAND
STRAND
           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
PIM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
N-LINKED (GLCNAC. . .) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULAR AMINOPEPTIDE COLLULAR AMINOPEPTIDE CAPEIN CAPEIN C.

CAPEIN PROTEIN C.

ENVELOPE GLYCOPROTEIN E1.

ENVELOPE GLYCOPROTEIN E2.

PROTEIN P7.

PROTEIN P7.

PROTEIN P7.

PROTEIN P82.

PROTEIN P82.

PROTEIN P82.

PROTEIN P82.

PROTEIN P84.

NONSTRUCTURAL PROTEIN NS4A.

NONSTRUCTURAL PROTEIN NS5A.

NONSTRUCTURAL PROTEIN NS5A.

NONSTRUCTURAL PROTEIN NS5A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (BY
(BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro: IPRO01409 HOV_NSSA.
InterPro: IPR001669 HOV_NSSA.
InterPro: IPR001669 HOV_NSSA.
InterPro: IPR001669 HOV_NSSA.
InterPro: IPR0016695 RNA_POL_DS_FS.
InterPro: IPR007094; RNA_POL_DS_FS.
InterPro: IPR007094; RNA_POL_DS_FS.
InterPro: IPR007094; RNA_POL_DS_FS.
InterPro: IPR007094; RNA_POL_PS_FS.
InterPro: IPR007094; RNA_POL_PS_FS.
InterPro: IPR007094; RNA_POL_PS_FS.
IPRAM: PF01539; HCV_NSS: I.
IPRAM: PF016001; HCV_NS3; I.
IPRAM: PF016001; HCV_NSA*; I.
IPRAM: PF01606; HCV_NSA*; I.
IPRAM: PF01601; HCV_NSA*; I.
IPRAM: PF01601; HCV_NSSA; I.
IPRAM: PF00166; HCV_NSSA; I.
IPRAM: IRRIPRIEME IRRIPRI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSFAC; TUALLY,
INTERPRO; IPRO01410; DEAD.
INTERPRO; IPR002522; HCV_capsid.
INTERPRO; IPR002521; HCV_core.
INTERPRO; IPR002521; HCV_core.
INTERPRO; IPR0025319; HCV_NS1.
INTERPRO; IPR004109; HCV_NS3.
INTERPRO; IPR004109; HCV_NS3.
R INTERPRO; IPR001490; HCV_NS4.
R INTERPRO; IPR001490; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                           EMBL; M67463; AAA45534.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1026
1657
1711
1972
2420
3011
369
1107
11107
11107
1237
196
209
                                                                                                                                                                                                                                                                                                                                                                                                      PIR, A36814, GNWVCH.
PDB; 1HEI; 25-NOV-98.
PDB; 1A1V; 16-FEB-99.
PDB; 1A1R; 17-JUN-98.
MEROPS; S29-001; ---
MEROPS; U39-001; ---
MEROPS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            192
384
847
1027
1027
1027
1083
1083
11165
11165
1216
1316
1316
1316
208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
```

```
CHAIN
                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery
  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                   53
                                                                                                                                                                                                                  01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last aenoctation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1): Protein P7: NonStructural protein NS2 (P21)
(EC 3.4.22.-1; Protesse/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P65); Nonstructural protein
NS5B (P65) (P70) (RNA-directed RNA POlymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate Taiwan) (HCV).
Viruses; SRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [RNA](N).
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. CONSIGNS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
15; Indels
                                                                                                                                                                                                  PRT; 3010 AA
10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                            174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M84754; -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
Matches 170; Conservative
                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PDB; 1N64; 25-FEB 03. PDB; 1NS3; 08-APR 98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001; ..
                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                    Hepacivirus
                                                                                                                                                                                                   POLG_HCVTW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS;
                                                                                                                                                                                                   g
                                                         à
                                                                        q
                                                                                            ò
                                                                                                              셤
                                                                                                                                 ò
                                                                                                                                                  d
```

```
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN BE (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
ROAD POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC,
N-LINKED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; Score 867.5; DB 1; Pred. No. 1.8e-74; 18; Mismatches 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATP (POTENTIAL).
DECH BOX.
                                                      InterPro; IPR002169; HV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR001095; RNA_DOL_DS_PS.
InterPro; IPR007095; RNA_DOL_DS_PS.
InterPro; IPR007094; RNA_DOL_PS_VIT.
IPRAM; PP01506; HCV_NS2; I.
IPRAM; PP01506; HCV_NS4a; I.
IPRAM; PP01506; HCV_NS4a; I.
IPRAM; PP01506; HCV_NS5a; I.
IPRAM; PP01806; HCV_NS1; I.
IPRAM; PP01806; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             84.1%;
79.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            327047
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       115
191
383
729
1006
1615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2013
3010
369
1083
IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 163; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
```

ï

us-09-965-594-16.rsp

IPR002522;

```
Best Local Similarity 78.4 Matches 160; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KKGSVVIVGRIN---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
     ODARAMA MARKA MARK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                  1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                  54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the Swiss Institute of Bioinformatics and the EMBL outstation
3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNOVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-000 (Rel. 42, Last annotation update)
67-000 (Rel. 42, Last annotation update)
67-00 (Grebs) (Grebs)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. CONSISTS OF TWO PROTEINS: PROFEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                       174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; D11168; BAA01943.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PDB; 1A1Q; 25-WAR-98.
PDB; 1JXP; 14-JAN-98.
NEROPS; S29.001; -.
NEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; A45573; A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MCBI_TaxID=31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVJT
000269:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          between
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                       ò
                                                                                                                                                  g
                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                       g
```

```
CARSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSJ ( POTENTIAL).
PROTEASE/HELICASE NSJ (POTENTIAL).
PROTEASE/HELICASE NSJ (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                             IPR002868; HCV_NS5a.
IPR002166; HCV_RdRP.
IPR007095; RNA_DOL_DS_PS.
IPR007094; RNA_DOL_PSVIr.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PF02907, HCV_NS3; 1.
PF01006, HCV_NS4s; 1.
PF01506, HCV_NS4b; 1.
PF0051, helicase_C; 1.
PF00598, Viral_RGRP; 1.
                                                                                                                              HCV_NS4a.
                                                                                                                                                                                                                                                                                                             Pfam: PF01543; HCV_capsid; 1. Pfam: PF01542; HCV_core; 1. Pfam: PF01543; HCV_core; 1. Pfam: PF01559; HCV_sns; 1. Pfam: PF01560; HCV_NS; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               326573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               645
2041
2077
2240
2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEXDC;
                                                               IPR002518;
                                                                                                IPR004109;
IPR000745;
IPR002519;
                                    PR002531;
                                                                                                                                                           IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SM00487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF01006;
Pfam; PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam; PF01506;
Pfam; PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                InterPro;
InterPro;
                                                                                                                                                                                                                                                               nterPro;
                                                                                                                                                                                                                                                                                             interPro;
```

Ä 53

9; Gaps

Indels

DB 1; Length 3010;

1.6e-73

83.1%;

20; Mismatches Score 857.5; Pred. No. 1.6

---LSGDIAYAQQIRGEEGCQEISQIGRDKNQVEGEVQIVSI

```
X. MEDLINE-98227846; PubMed-9568891;
YARY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.

YAR Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M., Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
T. Complex of N3 protease and NS4A peptide of BK strain hepatitis C.
T. VILUS: a 2.2 A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847 (1998).

- I- FUNCATION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 NAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 NAY PLAY A ROLE IN THE STACH bonds in the viral presursor polyprotein, commonly with Asp or Glu in the P6
presursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in Pl and Ser or Ala in Pl'.

C. CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
              RRGREILLGPADSIEGGGWRLLAPITAYAQQTRGLLGCIVTSLTGRDKNQVEGEVQVVST 1064
                                                                         1065 ATQSFLATCVNGVCWTVFHGAGSKTLAGPKGPITOMYTNVDQDLVGWHAPPGARSLTPCT 1124
                                                                                                    CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                         01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope diycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NSI); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate BK) (H2V).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                             54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; **Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.

MEDLINE-97015088; PubMed-8861916;
LOVE R.A., Parge H.E., Mickersham J.A., Hostomsky Z., Habuka N.,
MOGNAW E.W., Adachi T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEGUENCE OF 1487-1500.
MEDLINE-96235224; PubMed-8647104;
BOTOWSKi P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein Kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                 PRT; 3010 AA
                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               from human carriers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepacivirus.
NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              {RNA} (N).
                                                                                                                                                                                                                                     RESULT 5
POLG_HCVBK
ID POLG_HCVBK
AC P26663;
                                                                                                       114
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                             ò
                                             ò
                                                                         8
                                                                                                       ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DR INTERPROJ 3049.011. -.

BR INTERPROJ 198001410; DEAD.

BR INTERPRO; IPR001522; HCV_capsid.

BR INTERPRO; IPR001522; HCV_capsid.

BR INTERPRO; IPR0015219; HCV_core.

BR INTERPRO; IPR0015219; HCV_NS1.

BR INTERPRO; IPR0010518; HCV_NS2.

BR INTERPRO; IPR0010450; HCV_NS4.

BR INTERPRO; IPR001095; HRAL_POL_DS_PS.

BR INTERPRO; IPR001095; HRAL_POL_PS_PS.

BR INTERPRO; IPR001095; HCV_CORE; 1.

BR Pfam; PP01543; HCV_CORE; 1.

BR Pfam; PP01543; HCV_CORE; 1.

BR Pfam; PP01006; HCV_NS1; 1.

BR Pfam; PP01006; HCV_NS1; 1.

BR Pfam; PP01006; HCV_NS3; 1.

BR Pfam; PP01009; NIRAL_RARP; 1.

BR Pfam; PP01000; HCV_NS3; 1.

BR Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOEPTIDASE.
CELLULAR AMINOEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJRX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN BS/P2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/P2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Oore protein; Coat protein; Envelope protein; Holicase; ATP-binding;
Iransmembrane; Nonstructural protein; Hydrolase; Serine protease;
                            LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           115
191
191
729
1006
11615
2013
3010
369
11007
11165
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PDB: 1CSJ: 08-NOV-99.
PDB: 1GX5: 09-APR-02.
PDB: 1GX6: 10-APR-02.
PDB: 1QUV: 26-JUN-00.
PDB: 80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR; A38465; GNWVTC.
PDB; 1A10; 25-MAR-98
PDB; 1JXP; 14-JAN-98.
PDB; 1NS3; 08-APR-98
PDB; 1C2P; 15-NOV-00.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           116
192
384
730
1007
1616
1863
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1083
1107
1165
1230
1316
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3D-structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
```

```
CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3010 AA; 327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       82.7%; Score 853.5; DB 1;
77.9%; Pred. No. 3.9e-73;
tive 21; Mismatches 15;
                                                                            (GLCNAC.
                                                                                                                                                                                                                        (GLCNAC.
                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                  GICNAC.
                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                               (GLCNAC.
                                                                                                        (GLCNAC.
             (GLCNAC.
                                                                                                                    GLCNAC.
                                                                                                                                                         (GLCNAC.
                                                                                                                                             GLCNAC
                                                                                                                                                                                  GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                 645
2041
22077
2240
22529
22788
1047
1067
1063
1074
1074
1085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
tes 159; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1158
1162
1168
1172
1172
1187
1189
1203
                                                                                                                                                                                  114
                                                                                                        CARBOHYD
CARBOHYD
                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                       CARBOHYD
CARBOHYD
                                                                   CARBOHYD
                                                                              CARBOHYD
                                                                                                                                  CARBOHYD
                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
TURN
STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TURN
STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                     STRAND
STRAND
STRAND
TURN
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                           TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HELIX
STITE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                 01-A07-1992 (Rel. 23, Created)
01-A07-1992 (Rel. 23, Last sequence update)
01-A07-1992 (Rel. 24). Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
29-FEB-2003 (Rel. 41, Last annotation Edge)
34-21-37, Protease/helicase NS3 (P70) (Hepacivirin)
34-22-37, Nonstructural protein NS5A (P65); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
VIUUSES; SSRAM positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE-91192160; PubMed-1849488;
Rato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Rato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
Namotohon S., Shimotohon G. Labanese hepatitis C viral genome.";
"Molecular structure of the Japanese hepatitis C viral genome.";
"Molecular structure of the Japanese hepatitis C viral genome.";
"FEBS Lett. 280:325-328(1991).
"I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-:-CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral position, Cys or Thr in Pl and Ser or Ala in Pl'.
-:-CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                MEDIINE-91088550; PubMed-2175903; Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugimura T., Shimotohno K.; "Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis."; Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
3010 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL, D90208; BAA14233.1; -. PIR, A39253; GNWVCJ. HSSP; P26663; lJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS2
 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                 NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (RNA)(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
InterPro;
   POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
```

Φ

```
윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRINLSGD----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                          NONSTRUCTURAL PROTEIN N44 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                 CAPSID PROTEIN C (POTENTIAL).
MAIRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NON-STRUCTURAL PROTEIN NSI (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                             CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   82.7%; Score 853.5; DB 1; Length 3010; 77.0%; Pred. No. 3.9e-73; ive 23; Mismatches 15; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CRC64:
                                                                                                                                                                                                                                                                                                                                          N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AA993794F46DB185
                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                      POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
InterPro: IPR001650; Helicase_C.
InterPro: IPR007095; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_DS_PS.
InterPro: IPR007094; RNA_pol_DS_PS.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_NS1; 1.
Pfam; PF01560; HCV_NS2; 1.
Pfam; PF01000; HCV_NS3; 1.
Pfam; PF01000; HCV_NS4a; 1.
Pfam; PF01506; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5b; 1.
Pfam; PF01506; HCV_NS5b; 1.
Pfam; PF01506; HCV_NS5b; 1.
Pfam; PF01001; HCV_NS5b; 1.
Pfam; PF01001; HCV_NS5b; 1.
Pfam; PF01001; HCV_NS5b; 1.
Pfam; PF01001; HCV_NS5b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 645 645 N-LIN
2041 2041 N-LIN
2077 2077 N-LIN
2788 2248 N-LIN
3010 AA; 327017 MW; A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 77.0
Matches 157; Conservative
                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
CHAIN
  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEM. Outstarion the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE-92044440; PubMed-1658196;
MISCHORIDA A., MISCHORID M.;
MUCLECTIDE SEQUENCE of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.;
J. Gen. Virol. 72:2697-2704(1991).
J. Gen. Virol. 72:2697-2704(1991).
J. Gen. Virol. 72:2697-2704(1991).
J. Gen. Virol. 72:2697-2704(1991).
C. FUNCTION: THE SMALL PROTEIN NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PAZA A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PAZA A ROLE IN THE VIRAL RNA REPLICATION.
C. TARALTIC ACTIVITY: Hydrolysis of four peptide bonds in the P6 postition, Cys or Thr in Pl and Ser or Ala in Pl'.
C. CTALLTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                           01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3 4.22.-); Protease/helicase NS3 (P70) (Hepacivitin)
(EC 3 4.21.98); Nonstructural protein NS5A (P65); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P65) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J6) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flavivridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               {RNA}(N).
LIPOPROTEIN THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                         PRT; 3033 A.A.
1185 CTRGVARAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Interpro; IPR004109; HCV_NS3.
Interpro; IPR000445; HCV_NS4a.
Interpro; IPR001490; HCV_NS4a.
Interpro; IPR002866; HCV_NS5a.
Interpro; IPR00166; HCV_RGRP.
Interpro; IPR001665, Helicase_C.
Interpro; IPR001695; RNA_POl_DS_PS.
Interpro; IPR007094; RNA_POl_DS_PS.
Interpro; IPR007094; RNA_POl_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D00944; BAA00792.1; -.
                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JQ1303; JQ1303.
; P27958; IHEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HSSP; P27958; 1HE
MEROPS; S29.001;
                                                                                                                                                                                                                                                     POLG_HCVJ6
P26660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepaciv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEROPS
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRKGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R Pfam; PF01539; HCV_env; 1.

R Pfam; PF01589; HCV_NS1: 1.

R Pfam; PF01589; HCV_NS2: 1.

R Pfam; PF01580; HCV_NS3: 1.

R Pfam; PF01000; HCV_NS4: 1.

R Pfam; PF01000; HCV_NS5: 1.

R Pfam; PF01001; HcV_NS5: 1.

R Pfam; PF00271; Helicase_C: 1.

R Pfam; PF00271; Helicase_C: 1.

R Pfam; PF00389; Viral_RRP; 1.

R Probom; PD186602; HCV_NS1: 1.

R Probom; PD18602; HCV_NS1: 1.

R PROBOM; PROBOM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPERIDASE.
CAPSID PROTEIN (FOTEWILAL).
MATRIX PROTEIN (FOTEWILAL).
MAJOR ENVELOPE PROTEIN E (POTEWILAL).
NONSTRUCTURAL PROTEIN NS. (FOTEWILAL).
PROTEASE/HELICASE NS. (FOTEWILAL).
NONSTRUCTURAL PROTEIN NS.A (POTEWILAL).
NONSTRUCTURAL PROTEIN NS.A (POTEWILAL).
NONSTRUCTURAL PROTEIN NS.A (POTEWILAL).
ROMSTRUCTURAL PROTEIN NS.A (POTEWILAL).
RNA-DIRECTED RNA POLYMERASE (POTEWILAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 65.6%; Score 677; DB 1; Length 3033; 68.7%; Pred. No. 3e-56; ive 26; Mismatches 30; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC...)
(GLCNAC...)
(GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               329165 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        123; Comservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2038
2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     649 64
1091 109
2038 203
2811 281
3033 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 123; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2018
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

3033 AA

PRT;

STANDARD;

RESULT 8
POLG_HCVJ8
ID POLG_HCVJ8
AC P26661;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last sequence update)
03-REB-2003 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NSAA (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NSAA (P4); Nonstructural protein
NS5B (P66) (P70) (RMa-directed RNA polymerase) (EC 2.7.7.48)).
Hepatitis C virus (isolate HG-J8) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                           VITOLOGY 188:331-341(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in P1 and Ser or Ala in P1'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -:- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA.
                                                                                                                                                                                                                                                                                                                                                         *Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ); IPR007095; RNA_pol_DS_PS.
1PR007094; RNA_pol_DS_PS.
101543; HCV_core; 1.
10153; HCV_env; 1.
10159; HCV_env; 1.
10159; HCV_env; 1.
10159; HCV_env; 1.
102907; HCV_NS3; 1.
102907; HCV_NS4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4b; 1.
HCV_NS5a; 1.
Viral_RdRP; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                          WCBI_TaxID-11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01506;
PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (RNA)(N).
                                                                                                                                                                                                                             Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                              genotypes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        iga;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            79 IASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78
                                                                                                                                           CASID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR EWERLOPE PROTEIN (POTENTIAL).
MAJOR EWERLOPE PROTEIN (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
        CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC...) (POTE N-LINKED (GLCNAC...) (POTE GLCNAC...) (POTE GLCNAC...) (POTE GLCNAC...) (POTE GLCNAC...) (POTE M-LINKED (GLCNAC...) (POTE M-LINKED (GLCNAC...) (POTE M-LINKED (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HHOA_ARATH STANDARD; PRT; 321 AA.
09SEL7; 049507;
16-CCT-2001 (Rel. 40, Created)
16-CCT-2001 (Rel. 40, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Protease HHoA, chloroplast precursor (EC 3.4.21.-).
HHOA GRAG18370 OR F28312.30.
Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED GECNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; DB 1;
4.6e-56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL)
DECH ROY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. 4.6e.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2811
a: 330177 MW;
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  649
1091
2038
2359
2811
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                          TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HHOA ARATH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loca
Matches
                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                             CHAIN
  NAME OF A COLOR OF A C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
```

Arabidopsis thaliana (Mouse-ear cress). Eukaryota, Viridiplantae; Streptophyta, Embryophyta, Tracheophyta, Spermatophyta, Magnoliophyta; eudicotyledons; core eudicots, Rosidae;

```
We like the columnia of the co
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                 Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "The chloroplast lumen from Arabidopsis thaliana."; .
Submitted (JuL-201) to the SWISS-PROT date bank.
-1. SUBCELLULAR LOCATION: Chloroplast: within the thylakoid lumen.
-1. SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-1. CAUTION: Ref. 2 sequences differ from that shown due to erroneous gene model prediction. AT4G18370 and AT4G18375 were originally fused into a single gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320. Schubert M., Peterson U., Funk C., Haas B., Schroeder W.P., Kieselbach T.;
                                                                                                                              a homolog to the bacterial periplasmic protease HhoA."; Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases
Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nature 402:769-777(1999).
                                                                 SEQUENCE FROM N.A
                     NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   thaliana.
```

```
-1 - SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS
                                                                                                                                                                                                                                                               485 AA;
                                                                                                                                                                                                                                                                                                    Best Local Similarity
                                                                                                                                                         Complete proteome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TISSUE-Brain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAMP_HUMAN
Q13685;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  152
                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                     SIGNAL
                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                           DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 11
AAMP_HUMAN
                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               음
                                                                                                                                                                                                                                                                                                                                                                                                        - g
                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                        5
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                14;
           modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                     54 ATQTF-LATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC 112
                                                                                                                                                                                                                                                                                                                                                                                                 95 KTSPSVVYIEAIELPKTSSGDIL----TDEENGKIEGTGSGFVWDKLGHIVTNYHVIAKL 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                           203 VLGTSNDLRVGQSCFAI-----GNPYGYENTLTIGVVSGLGREIPSPNGKSISEAIQTDA 257
                                                                                                                                                                                                                                                                                                                                      2 KKKGSVVIVGRINL----SGDTAYAQQTRGEEGCQETSQTG--RDK--NQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                  ---SLLSPRPISYLK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Khalon B., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Atliach P., Bowman C., Cotton M.D., Fulii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.,
                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Science 281:375-388(1998).
                                                                                                                                                         Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide.
TRANSIT 1 26 CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema
  is in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              258 DINSGNSGGPLLDSYGHTIGV-NTATFTRKGSGMSSGVNFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  152 ----GSSGGPLLCPAGHAVGIFRAAVCTR---GVAKAVDF-IPVESLETTM 194
                                                                                                                                                                                                                                                                                                                26;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       'Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
                                                                                                                                                                                                                                                                                        Score 86; DB 1; Length 321;
Pred. No. 0.67;
 non-profit institutions as long as its content
                                                                                                                                                                                                                                                                                                              85; Indels
                                                                                                                                                                                                                                                       -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Y136_TREPA STANDARD; PRI; 485 AA. 083172; 16-0CT-2001 (Rel. 40, Created) 16-0CT-2001 (Rel. 40, Last sequence update) 16-0CT-2001 (Rel. 40, Last annotation update) 19-0thetical lipoprotein TP0136 precursor.
                                                                                                                                                                                           PROTEASE HHOA.
                                                                                                                                                                                                                                                                                                                28; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                   113 TCGSSDLYLVTRHADVIPVRRRGDSRG----
                                                                                                                                                                                 THYLAKOID
                                                                                                        InterPro; IPR001940; Proteasc2C.
InterPro; IPR001254; Scr_protease_Try
Pfam: PF00089; trypsin; 1.
                                                                                                                                                                                                       POLY-GLU
                                                                       EMBL; AL021710; CAA16717.1; ALT_SEQ
EMBL; AL161548; CAB78839.1; ALT_SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                             EMBL; AF114386; AAF24060.1; -
                                                                                                                                 Pfam; PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASES2C.
                                                                                                                                                                                                                                                                  34691 MW;
                                                                                                                                                                                                                                                                                        8.3%;
                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 26.89
Matches 62; Conservative
                                                                                                                                                                               71
321
87
145
186
                                                                                                                                                                                                                                                      40
321 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Treponema pallidum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                               MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Potential)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacteria;
                                                                                                                                                                                                                                           ACT_SITE
CONFLICT
                                                                                                                                                                                                                    ACT_SITE
                                                                                                                                                                                                                                 ACT_SITE
                                                                                                                                                                                                                                                                  SECTENCE
                                                                                                                                                                                 TRANSIT
                                                                                                                                                                                             CHAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7136_TREPA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license afreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVMSIATDGNTFVLACVP 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16 SGDTAYA-----QQTRGEEGCQETSQ----TGRDKNQVEGEVQIVSTATQTFLATCI- 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64 -NGVCWTVYHGAG----TRIIASPKGPVTQMYTNVDKDLVG-----WQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     109 GTGVYKHCVNGAGSSSTGTTASPSTETCSQHAT ---- LVGGTSKPFWLVPGGTGNNGNCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 C----GSSDLYLVTRHADVIP-----VRRRGDSRGSLLSPRPISYLK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               71;
                                                                                                                                                                                                                                                                                                                                      HYPOTHETICAL LIPOPROTEIN TP0136.
N-ACYL DEGLYCERIDE (POTENTIAL).
GLY/SER-RICH.
GLY/SER-RICH.
GLY/SER-RICH.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 485;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  48984 MW; C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                           Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 84.5; DB 1;
Pred. No. 1.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         225 DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGGG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----GSSGGPLLCPAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Anglo-associated migratory cell protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      452 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                            EMBL; AE001199; AAC65137.1; ALT_INIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    50; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (Human)
```

```
between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WIVYHGAGIRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  123 TRHADVIPVRRR---GDSRGS----LLSPRPISYLKGSSG--GPLLCPA----- 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                241 -----VLPDGKRAVVGYEDGTIRIWDLRQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      197 WMEWH-----PRAPVLLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzlaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
"Identification and characterization of Degp, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 163 -----GHAVGIFR----AAVCTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296 GSVDCQAKLVSATIGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Pred. No. 3;
13; Mismatches 47; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7.8%; Score 81; DB 1; Length 452;
25.3%; Pred. No. 3;
                                                                                                                                                                                                                                                                                                                   HEPARIN-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEGI_ARATH STANDARD; PRT; 437 AA.
022609, 09LK85;
16-0CT-2001 (Rel. 40, Last sequence update)
16-0CT-2003 (Rel. 40, Last annotation update)
28-FED-2003 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
DEGPI_OR DEGP_OR AT3G27925 OR K16N12.18.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DA1413D25EB236C0 CRC64;
                                                                                                                                                                   MIM, 603488; ...
MIM, 603488; ...
GO; GO:0008201; F:heparin binding activity; TAS.
InterPro; PRO01680; WD40.
Fig. PF00400; WD40; 8.
SWART; SM00320; WD40; 8.
PROSITE; PS00678; WD_REPEATS_1; 1.
PROSITE; PS0082; WD_REPEATS_2; 6.
PROSITE; PS50294; WD_REPEATS_2; 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                               POLY-GLU.
WD 1.
WD 3.
WD 4.
WD 5.
WD 6.
WD 6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           49015 MW;
                                                                                                                         EMBL; M95627; AAA68889.1; -. PIR; I39383; I39383.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 25.39 tes 42; Conservative
                                                                                                                                                                                                                                                                                                                  18
77
138
180
220
                                                                                                                                                       HGNC:18; AAMP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. STRAIN-cv. Columbia;
                                                                                                                                                                                                                                                                                                                                                             150
190
231
276
333
374
416
452 AA;
                                                                                                                                                                                                                                                                                                   Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID-3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                 DOMAIN
REPEAT
REPEAT
REPEAT
                                                                                                                                                            Genew;
                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                           REPEAT
                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   150 VPQGSSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---TRHADVIPVRRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         70 VYHGAGTRIIASPKGPVTQMY-----OA
                                                                                                                                                                                               ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                           STRAIN-cv. Columbia;
Kieselbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-!- FUNCTION: SERINE PROTESSE THAT IS REQUIRED AT HIGH TEMPERATURE.
MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
CAN DEGRADE BETA-CASEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Transit peptide; Chloroplast; Thylakoid.
           Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E:, Tabata S.;
*Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
                                                                                                                                                                                                                          -!- SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1497BlAB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHLOROPLAST (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    V -> I (IN REF. 2).
P -> S (IN REF. 2).
G -> R (IN REF. 2).
LL -> HF (IN REF. 2).
LL -> HF (IN REF. 2).
L -> V (IN REF. 2).
I -> V (IN REF. 2).
O -> E (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.6%; Score 78.5; DB 1;
                                                                                                                                                                                                                                                   -i- INDUCTION: By heat shock.
-i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-i- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROTEASE DO-LIKE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SERINE PROTEASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS: SOL. 279; --
InterPro: IPRO014478; PD2.
InterPro: IPR001940; Protease2C.
InterPro: IPR001254; Ser_protease_Try.
Pfam; PF00595; PD2; 1.
Pfam; PR00894; Programs, 1.
PRINTS; PR00834; PROFEASES2C.
SMART; SM00228; PD2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; AF028842; AAC39436.1; -.
EMBL: AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
MEDLINE-20363099; PubMed-10907853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  103 PQGSRSLTPCTCGSSDLYLV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 7.6%;
Best Local Similarity 25.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS50106; PDZ; 1.
                                                                    Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36
64
64
64
64
64
64
64
64
64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     321
421
171
201
280
23
                                                                                   [3]
SEQUENCE OF 104-118.
                                                                                                                                                                                                              O-PHENANTHROLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36
54
60
64
64
335
381
416
437 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        104
152
324
171
201
280
                                                       BAC clones."
                                                                                                                                                                                                                                          MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

```
Larsen F., Solheim J., Kristensen T., Kolsto A.B., Prydz H.; "A tight cluster of five unrelated human genes on chromosome 16q22.1.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hum. Mol. Genet. 2:1589-1595(1993).
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hum. Mol.
                                                                                                                                                                                                                                                                                                 ACT_SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SECUENCE
                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                   DOMAIN
                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                              DOMAIN
                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CTRL_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SÜBCELLULAR LOCATION: Type I membrane protein.

BYEKLOMENTAL STAGE: DFTECTABLE IN THE EPIBLAST OF OCCYTES AND THROUGHOUT EARLY MOUSE EMBRYO DEVELOPHENT. IN ADULT, EXPRESSION LOCALIZED IN GONADAL GERM CELLS.

SIMILARITY: Contains 2 protein tyrosine phosphatase domains.

SIMILARITY: Contains 10 fibronectin type III domains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)0 - protein
                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
NCBL_raxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBRYONIC STEM CELL PROTEIN TYROSINE
                                                 SYL-----KGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Lee K., Nichols J., Smith A.;
Mech. Dev. 61:213-215(1996).
-i. FUNCTION: MAY PLAY A ROLE IN THE MAINTENANCE OF PLURIPOTENCY.
DOWN REGULATED DURING DIFFERENTIATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tyrosine
                                                                        266 QDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIYSPSGASSGVGFSIPVDTV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00041; fn3; 7.

Pfam; PF00041; fn3; 7.

Pfam; PF00041; fn3; 7.

R PRINTS; PF00102; V_Phosphatase; 1.

R PRINTS; SM00060; FN3; 8.

SMART; SM000194; PTPC; 1.

DR PROSITE; PS000194; TYR_PHOSPHATASE_1; 1.

DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS50055; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 2.

Hydrolase; Transmembrane; Repeat; Signal; Glycoprotein.

POTENNIAL.

POTENNIAL.

POTENNIAL.
                                                                                                                                                                                                                                                15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Embryonic stem cell protein tyrosine phosphatase precursor
(EC 3.1.3.48) (ES cell phosphatase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE-Embryonic stem cells;
WEDLINE-97109513:
Lee K. Nichols J., Smith A. Smith A.
Tidentification of a developmentally regulated protein thosphatase in embryonic stem cells that is a marker of pluripotential epiblast and early mesoderm.";
Mech. Dev. 59:153-164(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EXTRACELLULAR (POTENTIAL)
POTENTIAL.
                                                                                                                                                                                                           PRT; 1705 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PHOSPHATASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MGD; MGI:108027; ESP.
InterPro; IPR003961; FN_III.
InterPro; IPR000387; TYR_phosphatase.
InterPro; IPR000242; Tyr_PP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; U36488; AAC52868.1; -. HSSP; P18052; 1YFO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     tyrosine + phosphate.
                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1077
                                                                                                                                                                                                                                                                                                                                                                                     musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19
1078
                                                                                                                                                                                                           PTPO_MOUSE
P70289;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
TRANSMEM
                                                 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ERRATUM.
                                                                                                                                                                                      PTPO_MOUSE
                                                                                                                                                              RESULT 13
                                                                                                                                                                                                                                  PARTIE E LA PRESENTA DE LA PRESENTA DEL PRESENTA DE LA PRESENTA DEL PRESENTA DE LA PRESENTA DEL PRESENTA DE LA PRESENTA DEL PRESENTA DE LA PRESENTA DE LA PRESENTA DEL PRESENTA DEL PRESENTA DE LA PRESENTA DE LA PRESENTA DEL PRESENTA 
                                                 ö
```

```
ж
;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----VVDKDL-VGW 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         442 TFCGLVPGAHYRV-----DIASSMGDITQSLTGYTSPLPPQSLEIISRNSPSDLTIGW 494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9 IVGRINLSGDTAYAQQTRGE----EGCQETSQTGR------DKNQVEGEVQIVSTATQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
FIBRONECTIN TYPE-III 1.
FIBRONECTIN TYPE-III 2.
FIBRONECTIN TYPE-III 3.
FIBRONECTIN TYPE-III 4.
FIBRONECTIN TYPE-III 6.
FIBRONECTIN TYPE-III 6.
FIBRONECTIN TYPE-III 7.
FIBRONECTIN TYPE-III 9.
FOOTEIN-TYROSINE PHOSPHATASE 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7.6%; Score 78; DB 1; Length 1705; 55%; Pred. No. 28; ve 12; Mismatches 53; Indels 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              101 -QAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSP 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    495 APAP-----GOMEGYKVTWHOD------GSORSP 517
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chymotrypsin-like protease CTRL-1 precursor (EC 3.4.21.-). CTRL OR CTRL1. Homo sapiens (Human).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2783755F15387D5B CRC64;
                                                                                                                                                                                                                                                                                                                                                           (GICNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                GLCNAC. . . .
                                                                                                                                                                                                                                                    N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58 FLATCINGVCWTVYHGAGTRTIASPKGPVTQMYT
                                                                                                                                                                                                                                        BY SIMILARIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-94093544; Pubmed-8268911;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     186795 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 25.5
hes 42; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                            905
1350
1350
1350
1350
1117
1117
1239
1239
1259
1259
1259
1259
1250
1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          982 90
1705 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CTRL_HUMAN
P40313;
```

MEDLINE-89073750; PubMed-3201749;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb.ch).
                                                                                                                                                                                                                                                                                                                                                                     ACTIVATION.

ACTIVATION PEPTIDE (POTENTIAL).

CHYMOTRYPSIN-LIKE PROTEASE CTRL-1.

CHARGE RELAY SYSTEM (BY SIMILARITY).

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.

BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 264;
                                                                                                                                                                                                                                                           SIMILARITY.
3F629F02FA6DDFB4 CRC64;
                                                                                                                                                                                       GO: GO: 0005615; C: extracellular space; TAS.
GO: GO: 0005605; C: extracellular space; TAS.
GO: GO: 0007586; P: edigestion; TAS.
GO: GO: 0006508; P: proteolygis and peptidolysis; TAS.
InterPro: IPR001314; Chymotrypsin.
InterPro: IPR001254; Ser_protease_Try.
                                                                                                           EMBL: X71874; CAA50710.1; EMBL; X71877; CAA50711.1; PIR; 138136; 138136. HSSP; p00763; 1DPO. MEROPS; SO1.256; Genew; HGNC: 2524; CTRL. MIW; 118888;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28002 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            210 ;
264 AA;
                                                                                                                                                                                                                                                                                                                                                                            PROPEP
CHAIN
ACT_SITE
ACT_SITE
ACT_SITE
DCSUBENTO
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
```

```
44 VEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAP 103
                                                                                                       118 MNNDVTLLKLASPAQYTTRISPVC------LASSNEALTEGLTCV---TTGWGRL 163
                                                                                                                                                                     104 QGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPI----SYLKGSSGG
                                    51; Indels 31; Gaps
                   4.3;
                                    18; Mismatches
7.4%; Score 76.5; 25.9%; Pred. No. 4.
                                                                                                                                                                                                                   157 PLLCPAGHA --- VGI 168
                                                                                                                                                                                                                                            11: :: :!!
217 PLVCQKGNTWVLIGI 231
                                    Conservative
                 Similarity
 Query Match
Best Local 8
                                    Matches
                                                                                                         g
                                                                                                                                              ò
                                                                                                                                                                                                                   ö
```

Squirrel monkey retrovirus (SMRV-H) (SMRV-HLB). Viruses; Retroid viruses; Retroviridae; Betaretrovirus.

SEQUENCE FROM N.A. NCBI_TaxID-11856;

01-MAY-1991 (Rel. 18, Created) 01-MAY-1991 (Rel. 18, Last sequence update) 28-FEB-2003 (Rel. 41, Last annotation update) PRT.

STANDARD;

VPRT_SMRVH P21407;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            93 VDKDLVGWQAPQGSRSLTPCTCGSSDLYLV---TRHADVIPVRRRGDSRGSLL----SPR 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         151 FHKPYRGASAP------GSSDVYWVQQISQQRPTLKLKLNGKLFSGILDTGADAT 199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42 NOVEGEVOIVSTATQTFLATCINGVCWTVYHGAGIRTIASPKG-------PVTQMYTN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                47; Gaps
                                                gene structure
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 323;
                                       Oda T., Ikeda S., Watanabe S., Hatsushika M., Akiyama K.,
Mitsunobu F.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  49; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY SIMILARITY.
5D6CEA38BA932786 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.4%; Score 76.5; D
23.3%; Pred. No. 5.5;
lve 16; Mismatches
                                                                                                                                                                                                                                                                                                                                      InterPro: IPR001995; Aspprotease_rtrv.
InterPro: IPR001969; Aspprotease_site.
InterPro: IPR001428; DeoxyUTPase.
InterPro: IPR001428; DeoxyUTPase.
InterPro: IPR001428; Gpatch.
Pfam: PF01585; GrPatch: 1.
Pfam: PF00777; rvp: 1.
Pfam: PF00777; rvp: 1.
PR081TE: PS00147; ASP_PROTEASE; 1.
PROSITE: PS00147; ASP_PROTEASE; 1.
PROSITE: PS0174; G_PATCH: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       146 PISYLKGSSGGPLLCPAGHAVGIFRA 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     200 VISYTHWPRNWPLITVATHLRGIGGA 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: August 30, 2003, 19:13:46
Job time : 10.7567 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     G-PATCH
                                                                                                                                                                                                                                                                               EMBL; M23385; AAA66452.1; ALT_INIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 323 AA; 35126 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 23.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                34; Conservative
                                                                                                                                                                                                                                                                                              PIR; B31827; PRLJHD.
HSSP; P06968; 1EUW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hydrolase; Aspartyl
                                                                                                                                                                                                                                                                                                                              MEROPS; A02.UPW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

hepatitis hepatitis

091rs1 091rt2 091rt2 091rt1 091rs9 091rs3 091rs3 091rs4 091rs7 091rs2 091rs2 091rs2 091rs2 091rs2 091rs2 091rs3 091rs2 091rs3 091rs3 091rs3 091rs3 091rs3 091rs3 091rs3 091rs3

hepatitis hepatitis hepatitis hepatitis

hepatitis hepatitis

hepatitis hepatitis hepatitis

ALIGNMENTS

```
SEQUENCE FROM N.A.
MEDIJNE-2032384; PubMed-10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Lai V.C., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lai V.C.H., Hong Z.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF268278; AAF82566.1; ...
HSSP: P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 4040 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Interpro: IPR004109; HCV_NS3.
Interpro: IPR002166; HCV_RdRP.
Interpro: IPR001056; Helicase_C.
Interpro: IPR001005; Myb_DNA_binding.
Interpro: IPR001005; RNA_bol_DS_PS.
Interpro: IPR007095; RNA_pol_DS_PS.
Interpro: IPR007094; RNA_pol_DS_PS.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF00998; Viral_RdRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR000280; CDvir_endptseP80
InterPro; IPR001410; DEAD.
 091R81
091R08
091R00
091RR0
091RR6
091R85
091R81
091R83
091R85
091R87
091R87
091R87
091R87
091R87
091R80
091R80
091R80
091R80
                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                           070817
Q91RT8
                                                                                                                                                                                                                                                                                 070818
091RR9
09DTE2
                                                                                                                                                                                                                                                                                                                                                                                                                                01-OCT-2000 (TTEMBLEEL 15, 01-OCT-2000 (TTEMBLEEL 15, 01-MAR-2003 (TTEMBLEEL 23,
 PRELIMINARY;
Genome polyprotein.
Mucosal disease virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pestivirus.
NCBI_TaxID-11099;
 MEROPS; S31.001
 872
870.5
870.5
869.5
869.8
869
869
869
868.5
868.5
                                                                                                                                                                                                                                                                                           866.5
866.5
866.5
091FH8
                                                                                                                                                                                                                                                                                                                                                                                     RESULT 1
 091FH8
                                                                                                                                                                                                                                                                                                                                                                                                            DORRAN REPARENCE OF THE PRESENT OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             091fh8 mucosal dis 036579 hepatitis c 081756 hepatitis c 09165 hepatitis c 0968 hepatitis c 03608 hepatitis c 099wx5 hepatitis c 091rr8 hepatitis c 091rr5 hepatitis c 091rr2 hepatitis c 091rr3 hepatitis c 091rr4 hepatitis c 091rr4 hepatitis c
                                                                                       (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,
                                                                                                                                             1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                            ; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         score greater than or equal to the score of the result bein and is derived by analysis of the total score distribution.
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                         Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                     830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                       summaries
                                                       OM protein - protein search, using sw model
                                                                           August 30, 2003, 19:00:22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091FH8
036579
081756
091FE5
099ELS8
003463
036608
099WX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1RR5
1RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                091RT9
091RR3
091RR4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1RT5
                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                          sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sp_vertebrate:*
                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1
                                                                                                                                                                                                                                                                                                                                                                                                                   sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                                                                        sp_archea:*
sp_bacteria:*
                                                                                                                                                                                                                                               Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_plant:*
sp_rodent:*
sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB
                                                                                                                                                                                                                                                                                                                              SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                               sp_fungi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4040
3011
2436
3011
3011
3011
3015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                                                                                                Sequence:
                                                                                                                                                                                                      Searched:
                                                                                                                                                                                                                                                                                                                                Database
                                                                               on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ş.
```

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q81756
Q81756;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AC DOTA NA WAY N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62 CINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone S.M., Rice C.M.; "Transmission of hepatitis C by intrahepatic inoculation with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          transcribed RNA.;
Science 277:570-574(1997).
-1-SUBGUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROYEN BY AND GIVCOPROTEINE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIVCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MANA (BY SIMILARITY).
EMBL; AFROD9606; AAB66324.1; -...
INTERTO: IPRO04610; DEAD.
InterPro: IPRO0410; DEAD.
InterPro: IPRO04522; HCV_cope.
                        SWART: SMO0487; DEXDC: 1.
SWART: SMO0480; HELICC: 1.
PROSITE; PSO0037; WIRE 1: 1.
PROSITE; PS50521; RDRP_VIRAL: 1.
PROSITE; PS50521; RDRP_VIRAL: 1.
PROSITE; PS00531; RNNSE_T2_2: 1.
ATP-blading; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase.
SEQUENCE 4040 AA; 453073 MW; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                               87.5%; Score 902.5; DB 12; Length 4040; 90.8%; Pred. No. 3e-81; Live 5; Mismatches 10; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-H77;
MEDLINE-97373636; PubMed-9228008;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_capsid.
HCV_core.
HCV_env.
HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RGRP.
PRINTS; PR00729; CDVENDOPTASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 90.8
Matches 177; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002531;
InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 2
036579
   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DRANGE DE LA DESTA DE LA DELLA DE LA DESTA DE LA DELLA DELLA
```

```
Pfam; PF01543; HCV_cave...,

DR Pfam; PF01542; HCV_core; 1.

DR Pfam; PF01564; HCV_core; 1.

DR Pfam; PF01569; HCV_NS1; 1.

DR Pfam; PF01569; HCV_NS2; 1.

DR Pfam; PF01009; HCV_NS4; 1.

DR Pfam; PF01001; HCV_NS4; 1.

DR Pfam; PF01001; HCV_NS4; 1.

DR Pfam; PF00209; VIral_RdFp; 1.

DR Pfam; PF00271; helicase.C; 1.

DR Pfam; PF00271; helicase.C; 1.

DR Pfam; PF00271; helicase.C; 1.

DR Pfam; PF00507; RDRP_POSITIVE; 1.

DR PROSITE; PS50507; RDRP_POSITIVE; 1.

DR PROSITE; PS50507; RDRP_POSITIVE; 1.

DR PROSITE; PS50501; RDRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1005 RRGQEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1125 CGSSDLYLVTRHADVIPVRRRGDGRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 ATQTFLATCINGVCWTVYHGAGTRIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genome polyprotein (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
Choo Q.-L., Richman K., Han J.;
Choo Q.-L., Richman K., Han J.;
The nucleotide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL; M32084; AAA45677.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .4e-80;
les 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 2436 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86.1%; Score 888.5;
84.3%; Pred. No. 5.4e
tive 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DR INCOMES AND THE TOTAL T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 84.3
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
```

m

```
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-colonel;
                          InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9ELS8;
Q9ELS8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              င်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01006; HCV_NS3: 1.

R Pfam; PF01001; HCV_NS4: 1.

R Pfam; PF01001; HCV_NS4: 1.

R Pfam; PF01001; HCV_NS5: 1.

R Pfam; PF001271; HCV_NS5: 1.

R Pfam; PF00271; HCV_NS5: 1.

R Pfam; PF00998; Viral_RGRP: 1.

R ProDom; PD186062; HCV_NS1: 1.

R SMART; SM00487; DEXCC: 1.

R PROSITE; PS50507; RDRP_POSITIVE; 1.

R PROSITE; PS50507; RDRP_VIMAL; 1.

R ATP-binding; Coat protein; Envelope protein; Glycoprotein; Hclicasc; W Hydrolase; Monstructural protein; Polyprotein; Mardolase; Monstructural protein; Polyprotein; Mardolase; Mardolase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Æ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDDINE-11262212; PubMed-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                   9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Genome polyprotein.

Hepatitis C virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   J. Geo. Virol. 82:1291-1297(2001).

-1- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL. AF211632: AAF81759.1:

HSSP: P27958: 1A1V.

InterPro: IPR003445: Cytc_heme_bind.
                                                                                                                                                                                                                                                                                                                                          Score 884.5; DB 12; Length 2436; Pred. No. 1e-79; 9; Mismatches 14; Indels 9;
                                                                                                                                                                                                                                                                                  2436 2436 24734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTRGVAKAVDFIPVENLETIMRSP 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS1.
                                                                                                                                                                                                                                                                                                                                       uuery Match 85.7%;
Best Local Similarity 84.3%;
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002518; H
IPR004109; H
IPR000745; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002521;
InterPro; IPR002519;
InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
InterPro;
                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                   NON TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9IFE5
Q9IFE5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 THE PARTY OF THE P
        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
```

```
1005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1065 AAQIFLATCINGYCWTVYHGAGTRIASPKGPVIQMYINVDQDLVGWPAPQGSRSLIPCI 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54 ATQIFLATCINGVCWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Desai S.W., Devare S., Yamaguchi J.;

"Hepatitis C Virus.";

"Hepatitis C Virus.";

Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.

-!- SUBGUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A ---

LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND WENAR (BY SIMILARITY).

EMBL, AF290978; AAG2099.1; ---

HSSP, P27988; 1HEI. InterPro; IPR0001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SMOUGHOUND CTROCHRONE_C; 1.
PROSITE; PS00190; RDRP_POSITIVE; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
ATP-binding; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ouery Match 85.7%; Score 884.5; DB 12; Length Best Local Similarity 84.3%; Pred. No. 1.4e-79; Matches 172; Conservative 9; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAR-2001 (TrEMBLrel. 16, Created)
1-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                  InterPro; IPR007095; NNA_POL_DS_C; Pfam; PP01543; NCV_capsid; 1. Pfam; PP01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01539; HCV_capsid; 1. Pfam; PF01539; HCV_cap; 1. Pfam; PF01539; HCV_cap; 1. Pfam; PF01001; HCV_NS3; 1. Pfam; PF01001; HCV_NS4; 1. Pfam; PF01001; HCV_NS4; 1. Pfam; PF01001; HCV_NS5a; 1. NF0100m; P1186062; HCV_NS1; 1. SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174 CTRGVAKAVDFIPVESLETTMRSP 197
IPR002868; HCV_NS5a.
IPR002166; HCV_RGRP.
IPR001650; Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
```

1;

53

Pfam;

```
"The 5'-terminal sequence of the hepatitis C virus genome.";
Jpn. J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 97E9052c0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001166; HCV_RdRP.
InterPro; IPR001650; Hellcase_C.
InterPro; IPR001050; RNA_DOL_DS_PS.
InterPro; IPR001094; RNA_DOL_DS_PS.
InterPro; IPR001094; RNA_DOL_PSv1r.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF015139; HCV_core; 1.
Pfam; PF01560: HCV_ENT; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410; DEAD.
InterPro; IPR001522; HCV capsid.
InterPro; IPR002521; HCV core.
InterPro; IPR002511; HCV core.
InterPro; IPR002511; HCV MS1.
InterPro; IPR002518; HCV MS2.
InterPro; IPR00140; HCV MS2.
InterPro; IPR00140; HCV MS2.
InterPro; IPR00140; HCV MS4.
InterPro; IPR00140; HCV MS4.
InterPro; IPR00186; HCV MS4.
InterPro; IPR00186; HCV MS4.
                                                                                                                                                                                                                                                                                                                                                              STRAIN-HC-J1;
MEDLINE-93117120; PubMed-1335573;
                                                                                                                          MEDLINE-92044440; PubMed-1658196;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pram; Pr01560; HCV_NS1; 1.
Pram; Pr01580; HCV_NS2; 1.
Pram; Pr01907; HCV_NS2; 1.
Pram; Pr01006; HCV_NS4s; 1.
Pram; Pr01001; HCV_NS4s; 1.
Pram; Pr01001; HCV_NS4s; 1.
Pram; Pr01505; HCV_NS5s; 1.
Pram; Pr00998; Viral_RdRP; 1.
ProDom; Pp186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             D10749; BAA01582.1; -. P27958; IHEI.
                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Inchauspe G.;
                                                                                                       STRAIN-HC-J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Okamoto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL;
SOW WE WANTED BY A STANDAY OF A STANDAY OF A STANDAY OF A STANDAY 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1125 CGSSDLYLVTRHADVIPVRRQGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C. virus...
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STKALN-HC-J1;
MEDLINE-91013116; PubMed-2170712;
Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          327107 MW; A6BECF5A3B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85.7%; Score 884.5; DB 1;
83.8%; Pred. No. 1.4e-79;
iive 11; Mismatches 13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                   InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_core.
InterPro; IPR002518; HCV_NS1.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR0014109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001566; HCV_NS5a.
InterPro; IPR001565; HSV_ARRP.
InterPro; IPR001055; HAILcase_C.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PSVIT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174 CIRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last seq
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PP01538; HCV_NS2; 1. Pfam; PP01538; HCV_NS2; 1. Pfam; PP01006; HCV_NS4; 1. Pfam; PP01001; HCV_NS4; 1. Pfam; PP01506; HCV_NS5; 1. Pfam; PP00971; hellcase_C; 1. Pfam; PP00998; viral_RdRP; 1. ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                      PF01543; HCV_capsid;
PF01542; HCV_core; 1.
PF01539; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity 83.8
Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY:
                                                                                                                                                                                                                                                                                                                                                                                                                     PF01560; HCV_NS1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART; SM
PROSITE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      003463
                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
Pfam;
```

a

a

õ

qq ð RESULT 6

```
Okamoto H., Okada S., Sugiyama Y., Kurai R., Iizuka H., Machida A., Miyakawa Y., Mayumi M.;
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J., Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Characterization and mapping of a B-cell immunogenic domain in characterization and mapping of a B-cell immunogenic domain in the petities C virus E2 glycoprotein using a yeast peptide library."; Virology 200:346-255(1954).

-i- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPPOPROTEIN ENTRE DATE DATE DATE OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Okamoto H., Kanai N., Mishiro S.; "Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-JI) with high homology to USA isolates."; Nucleic Acids Res. 20:6410-6410[1992].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-HC-J1;
MEDIINE-94174722; PubMed-7510436;
Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
```

```
IPR002868;
IPR002166;
                                                                                                                                                                                                                                                                                                                          Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               chimeras.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                              Ouery Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9PWX5
Q9PWX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q9PWX5
DAN KEN DAN KE
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ,,
                                                                                                                                                                                                   1005 RKGREILLGPADGMYSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                113
                                                                                                                                                                                                                                                                                                                       1065 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                            53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPPOPROTEINE.

-1 - SUBMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:

PROTEIN AND GLYCOPPOTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF EMBL. AFOILTS!

REMEL: AFOILTS! AAB67036.1;

RICEPTO: IPRO01410; DEAD.

RICEPTO: IPRO0252; HCV_core.

RICEPTO: IPRO0252; HCV_core.

RICEPTO: IPRO0251; HCV_core.

RICEPTO: IPRO0251; HCV_NS3.

RICEPTO: IPRO04109; HCV_NS3.

RICEPTO: IPRO0409; HCV_NS3.

RICEPTO: IPRO0409; HCV_NS3.

RICEPTO: IPRO0409; HCV_NS4.

RICEPTO: IPRO0409; HCV_NS4.

RICEPTO: IPRO04086; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
"Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a
                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY A
                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein.
Heparitis C virus strain H77.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus
                                                                                  6
                           DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JAN-1998 (TTEMBLrel. 05, Created)
01-JAN-1998 (TTEMBLrel. 05, Last sequence update)
01-MAR-2003 (TTEMBLrel. 23, Last annotation update)
                     Score 883.5; DB 1
Pred. No. 1.7e-79;
8; Mismatches 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1185 CTRGVAKAVDFIPVESLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR001650; Helicase_C.
IPR007095; RNA_pol_bS_PS.
IPR007094; RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-H77;
MEDLINE-97385173; PubMed-9238047;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_capsid; 1.
HCV_core; 1.
HCV_env; 1.
HCV_NSI; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS5a; 1.
helicase_C; 1
                     85.68;
84.38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS3; 1
                                                      Best Local Similarity 84.3
Matches 172; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01506;
PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          chimpanzee. "
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 7
036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DAY COOR ON BURN DE RESTRICT OF THE PARTY OF
                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                õ
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                  53
                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                  54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SMO4047; DEXDc; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 0B75E6B81CB5C198 CRC64;
                                                                                                                                                                                                                                                           9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

PROBLINE-9420395, PubMed-10489358;

Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;

"Hepatitis C virus: an infectious molecular clone of a second major manage (2a) and lack of viability of intertypic la and 2a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                 DB 12; Length 3011;
                                                                                                                                                                                                                                                           14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-2000 (TrEMBLrel. 13, Created)
1-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
                                                                                                                                                                                                                 85.4%; Score 881.5; DB 1 83.8%; Pred. No. 2.7e-79;
                                                                                                                                                                                                                                                        10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1185 CTRGVAKAVDFIPVENLGTTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 CTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Virology 262:250-263(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP; P27958; 1HEI.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                      Matches 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR001490;
```

```
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
NCBI_TaxID-11103;
                                                   P27958;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  091RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   091RR8
                     EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 291RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 ATQTFI.ATCINGVCWTVYHGAGTRT1ASPKGPVTQMYINVDKDLVGWQAPQGSRSI.TPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
       53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bukh J.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
-!- SUBGNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
-!- EUPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDINE-99420396; PubMed-10489358;
Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.;
"Hepatitis C virus: an infectious molecular clone of a second major
genotype (2a) and lack of viability of intertypic la and 2a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein.

Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepativirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 12; Length 3015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  85.4%; Score 881.5; DB 12; Length
83.8%; Pred. No. 2.7e-79;
iive 10; Mismatches 14; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1189 CTRGVAKAVDFIPVENLGTTMRSP 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-WAY-2000 (TrEMBLrel, 13, Created)
01-WAY-2000 (TrEMBLrel, 13, Last sequ
01-MAR-2003 (TrEMBLrel, 23, Last anno
IPR001650; Helicase_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 83.8 hes 171; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       chimeras.
 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q9PWU9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09PWU9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99PWU9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HERE REPRESENTED TO THE PROPERTY OF THE PROPER
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ద
```

```
54 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DMR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 processes (trempent).
Hepalitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85.4%; Score 881.5; DB 12; Length 3015;
83.8%; Pred. No. 2.7e-79;
tive 10; Mismatches 14; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50507; RDRP_VIRAL; 1.
ATP-Dinding; Coat protein; Envelope protein; Glycoprotein; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3015 AA; 328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR001212; HCV_capsid.
InterPro; IPR001212; HCV_capsid.
InterPro; IPR001213; HCV_capsid.
InterPro; IPR001213; HCV_capsid.
InterPro; IPR001213; HCV_capsid.
InterPro; IPR001218; HCV_NS1.
InterPro; IPR001419; HCV_NS3.
InterPro; IPR001445; HCV_NS4.
InterPro; IPR001446; HCV_NS4.
InterPro; IPR00145; HCV_NS4.
InterPro; IPR001216; HCV_NS3.
InterPro; IPR001516; HCV_NS3.
InterPro; IPR001505; HCV_NS3.
InterPro; IPR001505; HCV_NS1.
IPEAM; PF01543; HCV_capsid; IPP01543; HCV_Capsid; IPP01543; HCV_Capsid; IPP01543; HCV_NS1; IPP01540; HCV_NS2; IPP0160; HCV_NS3; IPP0160; HCV_NS4; IPP0160; HCV_NS4; IPP0160; HCV_NS4; IPP0160; HCV_NS4; IPP0160; HCV_NS4; IPP0160; HCV_NS4; IPP0160; HCV_NS1; IPP0160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 CTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
AF177039; AAF01181.1;
AF177037; AAF01179.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best_Local Similarity 83.8
Matches 171; Conservative
                                                                                                                                          InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEXDc;
                                                                                             THEI.
```

RR RA RE DR DR DR KW KW DR DR DR SO SO

a

ò

ö

a å

```
79 IASPRGPVTQMYTNVDRDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 TAYAQQTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.; "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IRN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865238, ARK545831;
Interpro: IPR004109; HCV_NS3.
Pfam: PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases EMBL/ RASS6241; ARS4656.1; -. InterPro; IRR04109; HCV_NS3. Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 181
181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  85.0%; Score 877; DB 12;
94.4%; Pred. No. 2.1e-80;
11ve 2; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                               Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                         01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23, NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 168; Conservative
                                                                                                                                                                                                                                                                        NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181 1
181 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-Pt.3U;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                              RESULT 12
Q91RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCOUNT REPORT R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYIVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64 IASPKGPVIQMYINVDRDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 TAYAQQTRGLLGCIITSLIGRDKNQVEGEYQIVSTAAQTFLATCINGVCWTVYHGAGTRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to Irw of the N33 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF865218; AAR5443.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                    STRAIN-Pt.1Y;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369235; AAK54560.1;
InterPro; IPR004109; HCV_NS3.
InterPro; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                   Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 879; DB 12; Length 1
Pred. No. 1.3e-80;
1; Mismatches 8; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                  181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q91RT5;
0-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MMR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                 Score 879; DB 12;
Pred. No. 1.3c-80;
1; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                   85.2%;
94.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tch 85.2%; al Similarity 94.9%; 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 94.77
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [1]
SEQUENCE FROM N.A.
                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-Pt.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
Matches 169
                                                                                                                                                                                                                                                                                                                                        NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                           NON_TER
```

g ò a

ö

æ

```
STRAIN-Pt.4B;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, ARS6240; ARX54565.1;
-InterPro; IPR004109; HCV_NS3:
Pfam; PF02907; HCV_NS3: I.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                     Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Search cor
Job time
     OX
REPRESENT OX
SO STATE OX
SO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64 IASPKGPVIQMYTWVDKDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                              79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                           19 TAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 TAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRI 63
                                                                                                                                          78
                                                                                                                                                                   4 TAYAQQTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAXQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                      GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    124 GGLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                       19 TAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-Pt.1C;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369214; AAK54639.1;
Interpro; IPR004109; HCV_NS3.
Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Gaps
                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DBC-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (targement).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      091RR3.
091RR3.
091RR3.
091RR3.
091RR3.
091DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2003 (TrEMBLrel. 23, Last sequence update)
01-NAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
NS3 protease (Fragment).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                   ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.0%; Score 877; DB 12; Length 181; 94.4%; Pred. No. 2.1e-80; tive 2; Mismatches 8; Indels
                         Length 181;
                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 181
181 AA; 19114 MW; BEIDOB542F014F86 CRC64;
                                                                                   .
8
                    Score 877; DB 12;
Pred. No. 2.1e-80;
2; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AA
                       85.08;
94.48;
                       Query Match
Best Local Similarity 94.49
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 85.07
Best Local Similarity 94.4*
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                      139
                                                                                                                                                                                                                                                                                                                                                                                                          124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091RT9
091RT9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q91RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 15
Q91RR3
AC Q91RR
AC Q91RR
DT 01-DE
DT 01-DE
DT 01-DE
DT N3 P
OS HEPAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                          ò
                                                                                                                                                                                        g
                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
ö
                                                                                                                                                                                                                        79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                 78
                                                                                                                                                                                                                                                                          139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                               19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT
                                                                                               Gaps
                                                                                            0;
                                                     84.9%; Score 876; DB 12; Length 181; larity 94.4%; Pred. No. 2.6e-80; Conservative 2; Mismatches 8; Indels
181 181
181 AA: 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                        Local Similarity
ses 168; Conserv
                   SEQUENCE
                                                       Query Match
```

completed: August 30, 2003, 19:18:19 e : 37.5921 secs

~

```
ALIGNMENTS
         em_htg_inv:*
em_htg_other:*
em_htg_mus:*
                                                                                 em_htgo_other:*
                               em_htg_pln:*
em_htg_rod:*
em_htg_mam:*
em_htg_wrt:*
                                                            em_sy:*
em_htgo_hum:*
   em_htg_hum:*
                                                                                                                                         Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                  9646
12980
12980
5360
5360
                                                                                                                                                                                                                                                                                                                                                                                            86.1
86.1
86.1
885.7
85.7
85.7
85.7
85.7
                                                                                                                                                 Score
                                                                                                                                                              9922.5
9908.5
9008.5
904.5
904.5
904.5
900.5
900.5
900.5
900.5
900.5
900.5
900.5
900.5
900.5
                                                                                                                                                                                                                                                                                              897.5
897.5
897.5
896.5
896
                                                                                                                                                                                                                                                                                                                                         893.5
893.5
                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                          Result
                                                                                                                                                                                                                          August 30, 2003, 19:18:33 ; Search time 2560.57 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                      US-09-965-594-16
1032
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                              5777422
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                             using frame_plus_p2n model
                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                               2888711 segs, 20454813386 residues
                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                            0.5
7.0
7.0
                                                                                                         Xgapop 10.0 , Ygapext (Ygapop 10.0 , Ygapext (Fgapop 6.0 , Fgapext Delop 6.0 , Delext
                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                           GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_ni:*

4: gb_ow:*

9b_ph:*

9b_ph:*

9b_ph:*

9b_ph:*

9b_ph:*

9b_pr:*

9b_pr:*

9b_pr:*
                                                                                                                                                                                                                                                                                                                                                                                                       gb_sy:#
gb_un:#
gb_vi:#
em_ba:#
em_fun:#
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_om:*
em_or:*
em_ov:*
em_pat:*
em_ph:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_in:*
                            - nucleic search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              em_ro:*
                                                                                                    BLOSUM62
                                                                       Title:
Perfect score:
                                                                                                    Scoring table:
                              OM protein
                                                                                                                                                                                                                                                                                                                  Database :
                                                                                       Sequence:
                                                                                                                                               Searched:
                                           Run on:
```

```
AR118692 Sequence
106440 Sequence 54
109329 Sequence 10
AR118696 Sequence
109331 Sequence
M32084 Hepatitis C
AR118703 Sequence
                                                                                                                                                                                               AR145264 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145265 Sequence
AR145267 Sequence
AR145267 Sequence
AR145269 Sequence
AR145269 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145273 Sequence
AR118686 Sequence
BD069982 Functiona
AR118686 Sequence
BD069985 Functiona
AR118686 Sequence
BD069985 Functiona
AR118686 Sequence
BD069985 Functiona
AR118686 Sequence
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                 SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR127810
BD081911
BD081911
BAR127809
BD081910
AR110828
BD069982
AR110831
BD069985
AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HPCPOLYP
AR118703
                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145253
AR145265
AR145265
AR145274
AR145272
AR145257
AR145250
AR145251
AR145251
AR145251
AR145251
AR145271
AR145271
AR145271
AR145271
AR145271
AR145271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          109328
AR118692
```

```
Unclassified
AR145264
AR145264.1
                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                               411 a
                     Unknown
                                 Unknown
                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                                   244
                                                                                                                                                                                                                                                                                                                                                                  304
                                                                                                                                                                                                                                                                                                                                                   82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
ACCESSION
VERSION
                                                                                                source
                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                              Pred. No.:
ACCESSION
VERSION
KEYWORDS
SOURCE
                                ORGANISM
                                                       AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 3
AR145268
                                                REFERENCE
                                                                               JOURNAL
                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAT 08-AUG-2001
         20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                 ValThrargHisAlaaspValIleProValArgArgArgAlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                          952
                                                                                                                                                                                                                                               532
                                                                                                                                                                                                                                                                                                                   592
                                                                                                                                                                                                                                                                                                                                                   652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                     41
                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                      21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                              GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                    AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                   CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                         pestivirus
         PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 linear
          linear
                                                               Unclassified.

1 (bases 1 to 12734)

Hong,Z., Lai,V.C.H. and Lau,J.Y.N.

Hong,Z. Lai,V.C.H. and Lau,J.Y.N.

Papelitis C virus protease-dependent chimeric patent: US 6326137-A 1 04-DEC-2001;
                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
         DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AR145264 1998 bp 1 Sequence 105 from patent US 6211338.
                                                                                                                                2803
     Sequence 1 from patent US 6326137.
AR179057.
AR179057.1 GI:20220612
                                                                                                                                                                                                                      US-09-965-594-16 (1-197) x AR179057 (1-12734)
                                                                                                                                                                                                     Gaps:
                                                                                                                    /organism-"unknown"
1 2604 c 3295 a
                                                                                                       Location/Qualifiers
                                                                                                                                                             8.58e-67
922.50
94.36%
92.31%
89.39%
                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                                                                4032 a
                                            .
Unknown.
                                                          Unknown
                                                                                                                                                     Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 893
                                                                                                                                                                                                                                                                                                                   533
                                                                                                                                                                                                                                                                                                                                                                                                  102
                                                                                                                                                                                                                                                                                                                                                                                                                                                 773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 833
                                                                                                                                                                                                                                      Ŋ
                                                                                                                                                                                                                                                     413
                                                                                                                                                                                                                                                                     22
                                                                                                                                                                                                                                                                                   173
                                                                                                                                                                                                                                                                                                    42
                                                                                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                                                    593
                                                                                                                                                                                                                                                                                                                                                                   82
                                                                                                                                                                                                                                                                                                                                                                                  653
                                                                                                                                                                                                                                                                                                                                                                                                                 713
                                                                                                                                                                                                                                                                                                                                                                                                                                 122
                                                                                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145264
LOCUS
DEFINITION
                                                                                                                 source
                 DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                               BASE COUNT
ORIGIN
                                                          ORGANISM
                                                                                AUTHORS
TITLE
JOURNAL
FEATURES
                                                                         REFERENCE
AR179057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                       Score:
         LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                     Š
                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                g
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                          243
                                                                                                                                                                                                                                                                                                                                                                               303
                                                                                                                                                                                                                                                                                                                                                                                                                    363
                                                                                                                                                                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                                                                                                                                                        423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        603
                                                                                                                                                                                                                                              183
                                                                                                                                                                                                                                                                                                                                                                                                  101
                                                                                                                                                                                                                                                                                     4
                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                             83
                                                                                                                                                                                                                                                                                               162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                                                                     22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT
                                              1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4 cofactor peptide
Patent: US 6211338-A 105 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTGGACTTTGTGCCGGTAGAGTCCATGGAAACTACTATGCGGTCTCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                           1998
170
16
7
3
                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR145268 1998 bp DV Sequence 109 from patent US 6211338. AR145268.1 GI:15107135
                                                                                                                         u
                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                              US-09-965-594-16 (1-197) x AR145264 (1-1998)
                                                                                                            /organism-"unknown"
595 c 569 q
GI:15107131
                                                                                                                                                           9.21e-67
911.50
94.90%
86.73%
88.32%
                                                                                                     .1998
```

```
Unknown.
Unclassified.
                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                42
                                                                                                                                                                                                                                            62
                                                                                                                                                                                                                                                                                    304
                                                                                                                                                                                                                                                                                                                                          424
                                                                                                                                                           'n
                                                      source
                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
ORGANISM
                                                                   BASE COUNT
                   AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                               AR145263
LOCUS
            REFERENCE
                                        JOURNAL
                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                          ò
                                                                                                                                                                      g
                                                                                                                                                                                     ò
                                                                                                                                                                                               q
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-AUG-2001
                                                                                                                                                                                                                             543
                                                                                                                                                                                                                                                                                                                                                                                               ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                       ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                        LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                            PAT
                         1 (bases 1 to 1998)
Malcolm.B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
protease and S14A cofactor peptide
Patent: US 6211338-A 109 03-APR-2001;
Location/Qualifiers
1.1998
                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                           1998
169
17
7
3
                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 103 from patent US 6211338. AR145262. GI:15107129
                                                                                                                                             Gaps:
                                                                                                                                                           US-09-965-594-16 (1-197) x AR145268 (1-1998)
                                                                         /organism-"unknown"
595 c 569 g
                                                                                                           1.64e-66
908.50
94.90%
86.22%
88.03%
    Unknown.
Unknown.
Unclassified.
                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                 411 a
                                                                                                      Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                                                                                  484
                                                                                                                                                                                                                                                          62
                                                                                                                                                                                                                                                                      244
                                                                                                                                                                                                                                                                                    82
                                                                                                                                                                                                                                                                                                  304
                                                                                                                                                                                                                                                                                                                             364
                                                                                                                                                                                                                                                                                                                                                       424
                                                                                                                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 4
AR145262
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                       Query Match:
DB:
                                                                   source
                                                                               BASE COUNT
ORIGIN
                                                                                                     Alignment :
Pred. No.:
              ORGANISM
                                 AUTHORS
TITLE
                                                     JOURNAL
FEATURES
                           REFERENCE
KEYWORDS
                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                  δ
                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                        οy
                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                         ò
                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                    ò
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                      243
                                                                                                                                                                                                                                                                                                                                                                                                                              303
                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                  64 GGITCIGIIGIIAIIGIIGGIAGAAIIAIIIIAACIGGIAGIAGIAGIATCACGGCCIAC
                                                                                                                                                                                                                                                                                                                                                       AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                        CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAlaSer
1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NSAR cofactor peptide
Patent: US 6211338-A 103 03 APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 104 from patent US 6211338. AR145263
                                                                                               424
                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                  US-09-965-594-16 (1-197) x AR145262 (1-1998)
                                                                                   /organism="unknown"
596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GI:15107130
                                                                                                                                              1.98e-66
907.50
94.39%
86.73%
87.94%
                                                                        1998
                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR145263.1
                                                                                                410 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
```

Unclassified

Unknown

```
Percent Similarity:
Best Local Similarity:
                                                 120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR145266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown
                                                                        Alignment Scores:
                                                                                                            Query Match:
DB:
                                  source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR145266
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
TITLE
                                                 BASE COUNT
                  JOURNAL
                          FEATURES
                                                         DRIGIN
                                                                                       Score:
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                 셤
                                                                                                                                                                                 ò
                                                                                                                                                                                               g
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                   ò
                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                        243
                                                                                                                                                                                                                                                                    LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                               21
                                                                                                                                                                                                 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                              81
                                                                                                                                                                  GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlafyr
                                                                                                                                                                                                                                                              CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 621338-A 104 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unclassified.
1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                             1998
170
15
8
8
                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NO
DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 95 from patent US 6211338.
AR145254
AR145254.1 GI:15107121
                                                               424
                                                                                                                                                  (1-1998)
                                                       /organism="unknown"
596 c 568 g
                                                                б
                                                                                                                                                   US-09-965-594-16 (1-197) x AR145263
                                                                                            1.98e-66
907.50
94.39%
86.73%
87.94%
                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                               410 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unknown
                                                                                       Scores:
                                                                                                                                                                                                                                                                                                                                                                                                      484
                                                                                                                                                                                                                                                                                                                                                                                                                                                    182
                                                                                                                                                                  S
                                                                                                                                                                                                55
                                                                                                                                                                                                               124
                                                                                                                                                                                                                                                                                            82
                                                                                                                                                                                                                                                                                                          304
                                                                                                                                                                                                                                                                                                                                                                      424
                                                                                                                                                                                                                                                                                                                                                                                      142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   604
                                                                                                                                                                                                                                                             62
                                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 6
AR145254
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                 sonrce
                                                                                       Alignment S
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
AUTHORS
                                                               BASE COUNT
REFERENCE
AUTHORS
TITLE
                                JOURNAL
FEATURES
                                                                       ORIGIN
                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                      염
                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                g
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                        ç
                                                                                                                                                                  à
                                                                                                                                                                                                ö
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                          64 GGTTCTGTTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                        183
                                                                                                                                                                                                                                                                                                                                                                                          142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                             GlySerValVallleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                               42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                             62 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 (bases 1 to 1998)
Malachim, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
  hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                        651
169
17
13
                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Z
N
O
  oŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8 bp DN
6211338.
Single-chain recombinant complexes oprotease and NS4A cofactor peptide Patent: US 6211338-A 95 03-APR-2001;
                                                                         144
                                                                                                                                                                                                         (1-651)
                                     Location/Qualifiers
                                                          /organism="unknown"
1 187 c 200 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 107 from patent US
AR145266
                                                                                                                                                                                                          US-09-965-594-16 (1-197) x AR145254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ARI45266.1 GI:15107133
                                                                                                                     1.03e-66
904.50
94.87%
86.67%
87.65%
                                                 .651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unclassified
```

```
AR145261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown
                                                                    Percent Similarity
                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                 122
                                                                                                                                                                                                                                                                                                                                                                                                                                 182
                                                                                                                                                                                                                                                                82
                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                                                 424
     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                   BASE COUNT
ORIGIN
                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
AR145261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
                                                            Score:
                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                             ò
                                                                                                                                            g
                                                                                                                                                              ò
                                                                                                                                                                            qq
                                                                                                                                                                                          õ
                                                                                                                                                                                                            a
                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT 08-AUG-2001
                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                         AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                              243
                                                                                                                                                                                                                                                                                                                                                                                                                                 603
                                                                                                                                                                                            124 TCCCAACAGAGGGGGGGGCTACTTGGTTGCAAGATCACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                                                             41
                                                                                                                                                                                                             42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                             ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                      TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                                                                                                                                           22 AlaGlnGlnThrArgGlyGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unclassified.

I (bases 1 to 1998)

Malcolm.B.A., Taremi.S.Shanc., Weber.P.C. and Yao,N.
Malcolm.B.A., Taremi.S.Shanc., Weber.P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 108 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 108 from patent US 6211338. AR145267. GI:15107134
 03-APR-2001
                                  424
                                                                                                           Gaps:
                                                                                                                            US-09-965-594-16 (1-197) x AR145266 (1-1998)
Patent: US 6211338-A 107 03-2
Location/Qualifiers
1. .1998
                        /organism-"unknown"
596 c 568 g
                                                                  3.53e-66
904.50
94.39%
86.22%
87.65%
                                                                                  Percent Similarity:
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
                                                           Alignment Scores
                                                                                                                                                                                                                             184
                                                                                                                                                                                                                                              62
                                                                                                                                                                                                                                                               244
                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                              304
                                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                               364
                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                124
                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                484
                                                                                                                                                                                                                                                                                                                                                                                                                162
                                                                                                    Query Match:
                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 8
AR145267
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
JOURNAL
FEATURES
                                                                   Pred. No.:
                                  BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                          ORIGIN
                                                                           Score:
                                                                                                            DB:
                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                            ò
                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                              đ
                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                            ò
                                                                                                                                                                                           셤
                                                                                                                                                                                                            á
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                            셤
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                           101
                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                     243
                                                                                                                                                                                                                                                                                                                                                                                304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG 363
                                                                                                                                                                                                                  4
                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                          AlabroGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                              42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                          5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide
Patent: US 621138-A 102 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                        1998
169
16
8
3
                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 bp D
6211338.
             424
                                                                                                                     Gaps:
                                                                                                                                          US-09-965-594-16 (1-197) x AR145267 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1998 Sequence 102 from patent US 6
/organism="unknown"
596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /organism-"unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ARI45261.1 GI:15107128
                                                      3.53e-66
904.50
94.39%
86.22%
87.65%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unclassified,
                                                                                          Best Local Similarity:
Query Match:
DB:
            ĸ
```

Alignment Scores: Pred. No.: Score: Percent Similarit: Best Local Similari Query Match:								
	es: rity: ilarity:	4.28e-66 903.50 93.88% 86.73% 87.55%	Length: Matches: Conservati Mismatches Indels: Gaps:	h: es: rvative: tches: s:	1998 170 14 3			
-+65-596-60-SN	(1-197) x AR145261 ((1-1998)					
0y 5 0	GlyServal	ValileValGlyAz	gileAsn	LeuSerGly	ds	-ThrAlaTyr	21	
Db 64	GGTTCTGTT(SARTATT	TTATCTGGT	AGTGGTAGTA 1	CACGGCCTAC	123	
Oy 22 Db 124	AlaGlnGln::::	AladindinThrArgGlyGluGlyGysGlnGluThrSerGlnThrGlyArgAspLys :::	luglyCys	GlnGluThr:	SerGlnThrGl	yargaspLys 	41	
~	AsnGlnVal(InIleval	SerThrAla1	ChrGlnThrP!	neLeuAlaThr	9	
Db 184	AACCAGGTC	GAGGGAGAGGTTC	GGTGGTT	TCCACCGCA	CACAATCCT	CCTGGCGACC	243	
Qy 62	CyslleAsn(::: TGCGTCAACC	CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 	rValTyr	HisglyAla(SlyThrArgT}	nrileAlaSer	81	
2 2	Droftwee lw	over training		Acres Acres	To A con Law	200000000000000000000000000000000000000	5 5	
30	FICEYSOLY 	FILINGSOLYTOGATION CLUMCLY THE ABJUANCE AND ASSAULATION TO THE CANAGEGE COCACCACCACCACCACCACCACCACCACCACCACCACCA	GTACACT	AATGTGGACC		TCGCTGCAG	363	
Oy 102	AlaProGln	GlySerArgSerLe	uThrPro	CysThrCysC	SlySerSerAs	spleuTyrLeu	121	
Db 364	000000000	GGGCCCCCGGGGGCGGTTCCTTGACACCATGCACCTGTGGGCAGCTCAGACCTTTACTT	GACACCA	TGCACCTGT	SGCAGCTCAG	ACCTTTACTTG	423	
Oy 122 Db 424	ValThrargi GTCACGAGAG	ValThrarghisalaaspValIleProValargargArgGlyaspSerArgGlySerLei 	leProval	ArgArgArg(cgccgccgc	31yAspSerA1 	rgGlySerLeu 	141	
0v 142	LenSerPro	Argprofleder#1	27.[110.]77	Gluserser	.lordv[5v].	Orden	161	
4 8 4	CICICCCC	Defined From great for the features of the fill	CTTGAAG	GGCTCTTCGC	STECTORY STORY	edbedcyskio 	543	
Qy 162	AlaGlyHis	AlaValGlyIleP	neArgAla	Alavalcys	rhrArgGlyVa	alAlaLysAla	181	
Db 544	TCGGGGCAC	:::	rccgggcr	GCCGTATGC		TGCGAAGGCG	603	
Qy 182 Db 604	ValAspPhe GTGGACTIT	ValaspPhelleProValGluSerLeuGluThrThrMetArgSerPro 	erLeuGlu ::: CATGGAA	ThrfhrMeta 		197 651		
RESULT 10 AR145269 LOCUS LOCUS LOCUS LOCUS LOCUS ACCESSION AR VERSION AR KEYWORDS CORGANISM UN REFERENCE LOCUS AUTHORS TITLE SITILE SITILE SITILE SOURCE SOURCE COURT ORIGIN	AR145269 Sequence 11: AR145269 AR145269.1. Unknown. Unknown. Unclassiiee 1 (bases 1 Halcolm.B.A Single-chai protease an Patent: US 1 (1)	O from patent GI:15107136 d. to 2016) Taremi,S.S n. Taremi,S.S no recombinant d NS4A cofact 6211338-A 110 cation/Qualif 2016 rganism-"unkn rganism-"unkn	2016 bp US 6211 hane., W complex or pepti 03-APR- iers	A. C. hep	linear and Yao,N atitis C vi	PAT 08-AUG-	-2001	

```
Pestivirus type 1
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Pestivirus.
1 (bases 1 to 12734)
1 (bases 1 to 12734)
1 (bases 2 to 12734)
1 (bases 2 to 12734)
1 (bases 3 to 12734)
2001s, R.O., Hong, Z. and Lau, J.Y.
Generation and characterization of a hepatitis C virus NS3
protease dependent bovine viral diarrhea virus
20323484
                                                                                                                                                                                                                                                                                                                                                                                                    VRL 12-JUL-2000
                                                                          561
                                                                                                                                                                                                                                                                                                                       201
                                                                                                                                                                   142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                         41
                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                    RNA
                                                                                                                                                                                                                                                                                                                                                                                               Pestivirus type 1, complete genome.
                                                            US-09-965-594-16 (1-197) x AR145269 (1-2016)
                                                                                                                                                                                                                                                                                                                                                                                                                         AF268278.1 GI:9049956
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 (bases 1 to 12734)
Lai,V.C.H. and Hong,Z.
Direct Submission
       4.32e-66
903.50
93.88%
86.73%
87.55%
                      Percent Similarity:
Best Local Similarity:
Alignment Scores:
Pred. No.:
                                                                                                                                                                                  262
                                                                                                                                                                    62
                                                                                                                                                                                                 82
                                                                                                                                                                                                               322
                                                                                                                                                                                                                                                                          442
                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                   RESULT 11
AF268278
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                                                                                                                                                                à a à
                                                                                                                                                                                                           g & g
                                                                                                                                                                                                                                                       oy
oy
                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                  q
                                                                           g
                                                                                                                                D
D
```

JOURNAL

5'UTR CDS

```
LITKHWYRNETVASWYGEEEVYGMPKIMTITRASTLEKSRHCTICTVCEGREWKGGTC
PKOCRHGKPITCASKLADFEERHYKRIFIREGNEGAKGRCGKTGRRTRPEEMDREPKSAR
YCAECNRLHPAEEGDFWAESSMLGLKITYFALMOGKYYDITEWAGCORVGISPDTHRY
YCAECNRLHPAEEGDFWAESSMLGLKITYFALMOGKYYDITEWAGCORVGISPDTHRY
PCHISFGSRUPFROETVGRYOYTARGOLE-LANLPVLATKVKHAWYGRIGEEIGNLEHL
GWILRGPAVCKRITEHEKCHINILDKLTAFFGIMPRGTTPRAPVRFPTSLLKVRRGLE
TGARATHQGGISSYDHYTAGKDLLVODSMGRPRVVCORNRLTPETEGGETEGVAFDSCPD
GARCYVLAPEANN ISGSKAVHLOKTGEFTCYTASGTPAFFDLKNLKGWSGLPIFE
ASSGRVVGRYGKVGRNEESKPTKIMSCIQTVSKNTADLTEWVKKITSMNRGDFKOITLA
TGARATTELPKAVIEEIGHRKYUVLIPLRAABSTYTOYMRLHPSISCULRGDKE
GDWATGITYASYTGYGRYCHPIEERIAPFWYRGIFFCINGKIHFSES
IRVVAMTAPPAGSYTTTGGKHPIEERIAPFWHKGEDLGSOFLDIAGLKIPVBSKOM
LVFVPTRNMAVEVAKKLKAKGYNSGYYSGEDPANLRVVTSOSPYVIVATNAIESGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARKLGEDVPYY IYATEDEDLAVDLLGLDWPDPGNQQVVETGKALKQVTGLSSAENALL VALEGYVGYQALSKRIFPMITDIYTIEDQRLEDTHHQYAPPAIKTDGTETELKELAS GDVEKINGGI SDYAAGGLEFVKSQAEKIKIAPLFKENAEAAKGYVQKFIDSLIENKEE I IRYGLMGTHTALKSIAARLGHETAFATLULKWLAFGGESVSDHYKQAAVDILVYYYY WHKFSFFGDSETQQGGRRFVASLFISALATTYKTWNYHNLSKVVEPALAYLDYATSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LPDLDTVIDTGLKCEKRVRVSSKIPFIVTGLKRMAVIVGEQAORRGRVGRVKPGRYR
SQETATGSKDYHYDLLQAQRYGIEDGINVTKSFREMNYDWSLYEEDSLLITQLEILNN
LLISEDLPAAVKNIMARTDHPEPIQLAYNSYEVQVPVLFPKIRNGEVTDTYENYSFLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YILDLIYGLHKQINRGLKKMVLGWAPAPFSCDWTPSDÈRIRLPTDNYLRVETRCPCGY
EMRAFKNYGGKLTKYEESGPFLCRNRPGRGPYNYRYTKYYDDNLREIKPVAKLEGQYE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AYTFVNEDVGTIKPVLGERVIPDPVVDINLQPEVQVDTSEVGITILGRETLMTTGVTP
VLEKVEPDASDNONSVKIGLDEGNYPGPGIQTHTLITEEIHNRDARPFIMILGSRNSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NRAKTARNINLYTGNDPREIRDLMAAGRMLYVALRDVDPELSEMYDFRGTFLDREALE
ALSLGQPKPKQVTKEAVRNLIEQKKDVEIPNWFASDDPVFLEVALKNDKYYLVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KSNRGHMASAYQLAQGNWEPLGCGVHLGT I PARRYK I HPYEAYLKLKDF I EEEEKKPR
VKDTV I REHNKWI LKK I RPQGNLNT KKMLNPGKLSEQLDREGRKRN I YNHQIGT I MSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RGKFWTTLLNGPAFOWCPIGWTGTVSCTSFNADTLATTVVRTYRRSKPFPHRQGTT

RKKTAELLHUGTAFVCPEDQLIFKTGGSTSCKWGCTGPRESECHPTPTGCKCL

ENETGYRLUDSTSCHREGWALVPGGTLKCK IGKTTVQVIAADTKLEBHRCRPYEIIS

EGPVEKTACTFNYTKTLKNKYFEPRDSYFQQYMLKGFYQVIAADTKLEPHRDYFAEIS

LYTLLRGGRYVLMTLYTWTLSBCRAGIQYGSGEVVMADKLETHNDFFAESI

LYTLLRESYKKWYLLLYHILWYHPTKSVYTYLLAGIGDVYKADSGGGYLGKTDLCFT

TVVLLVIGLIIARRDPTIVPLUYTHMALRYTELTHOPGVDIAVAWTITTLLAVSYYDD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LGVGA IAAHNA IESSEOKRTILAKVFVKNFILOQAATDELVKENPEKIINALFEAVOTI
GNPLRLIYHLYGVYYKGWEAKELSERTAGRNLFTLIMFEAFELLGMDSQGKIRNLSGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ERDCATITKNTVQFLKMKKGCAFTYDLTISNLTRLIELVHRNNLEEKEIPTATVTTWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /product="polyprotein"
/protein_id="AAF82566.1"
/bc.kref="5g1:90458651"
/translation="MELNYPEGSGSVVIVGRIVLSGSGSITACAQOTRGLLGCKITSL
/translation="MELNYPEGSGSVVIVGRIVLSGSGSITACAQOTRGLLGCKITSL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GKTKSKNTODGLYHNKNKPOESRKKLERALLAWAIIAIVLFOVTWGENITOWNLODNG
TEGIORAMPROKOVNSCHGINPEKICTOVPSHLATDIELKITHGWODSEKINYTOCR
LORHGWNKHOWCNWYN IEPWILVMNRTOWNTEGOPPRECAYTCRYDBASDLWYVYTOR
ROSPTPLTGCKKGKNFSFAGILANGPCNFEIAASDVLFKEHENISMFODTTLYLVDGL
TYSLECARGOTAKLTYWLGKOLGILGKKLENKSKTWYGAYAASPYCDYDRKIGYIWYT
KNCTPACLPKNTKIYGGKYCTOMAEDGKILHEWGGHLSEVLLLSLVALLSEPAPETASV
KNCTPACLPKNTKIYGGKFOTNAEDGKILHEWGGHLSEVLLLSLVALLSEPAPETASV
MYLILHFSIPOSHVDVMDCDKTQLNLTVELLTADVIPGSVWNLGKWVCIRPNWWPYET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YFRYKKWLOCILSLVSGVFLIRSLIYLGRIEMPEVTIPWRPLTLILLYLISTTIVTR
KVDYAGLLOCVPILLLYTUMADETTLILLILTYELKYLYKKTRYDIERSHUGG
IDYRRVDSIYDVDESGEGVYLEPSROKAQGENSILLPLIKATIISCVSSKWOLIYMSI
LTLDPMYYMHRKVIEEISGGTNIISRLVAALIELMSWEEEESKGLKKFYLLSGRLRN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HYYKGVTAKIDYSKCKMLLATDKWEVEHGVITRLAKRYTGVGFNGAYLGDEPNHRALV
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research
Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
SYSDTREEGATKKKTQKPDRLERGKMKIVPKESEKDSKTRPPDATIVVEGVKYQVRKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TVVLAFEEVSGVVKLVLRALRDLTRIWNAATTTAFLICLVKIVRGGMVOGILWLLIIT
GVQGHLDCKPEFSYAIAKDERIGQLGAEGLTTTWKEYSPGMKLEDTWVIAWCEDGKLM
YLORCTRETRYLAILHTRALPTSVVFKKLFDGRRQEDVVEMNDNFEFGLCPCDAKPIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LKMFTPTRLESVVILSTTIYKTYLSIRKGKSDGLLGTGISAAMEILSQNPVSVGISVM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VKDQAKALGATDQTRIIKEVGSRTYAMKLSSWFLQASNKQMSLTPLFEELLLRCPPAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRPISYLKG
                                                                                                                                                                                        /organism="Pestivirus type 1"
/mol_type="genomic RNA"
/db_xref="taxon:11099"
1 385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /codon_start-1
                                                                                                                                                                                                                                                                                                                                                                                                            .12508
```

```
THE METER TO THE METER THE METER THE METER TO THE METER TO THE METER THE THE METER THE METER THE METER THE METER THE METER THE METER THE THE METER THE METER THE METER THE METER THE METER THE METER THE THE METER THE METER THE METER THE METER THE METER THE METER THE THE METER THE METER THE METER THE METER THE METER THE METER THE THE METER THE METER THE METER THE METER THE METER THE METER THE THE METER THE METER THE METER THE METER THE METER THE METER THE THE METER THE THE METER THE METER THE METER THE METER THE METER THE METER THE M
YGEVTWEQLEAGINRGAAGFLEKKNIGEVLDSEKHLVEQLVRDLKAGRKIKYYETAI
PKNEKRDVSDDWQAGDLVVEKRPRVIQYPEAKIRLAIIKVMYNWVKQOPVVIPGYEGK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          593 TGCATCAATGGGGTATGCTGGACTGTCTACCACGGGGCGGGAACGAGGACCATCGCATCA 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 997
                                                                                                                                                                                                                                             1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AND
D
                                                                                                                                                                                                                                           ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AR145258 651 bp I
Sequence 99 from patent US 6211338.
AR145258
                                                                                                                                                                                                                                             2802
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-16 (1-197) x AF268278 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                3293 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GI:15107125
                                                                                                                                                                                                                                                                                                                                       3.99e-65
902.50
93.33%
90.77%
                                                                                                                                                                                                                                           2608 €
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145258.1
                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown.
Unknown.
                                                                                                                                                                                                                                                4030
                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         713
                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                    3'UTR
BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local
                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR145258
                                                                                                                                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
```

```
g
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                οy
                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                ò
                                                                                                                                                           셤
                                                                                                                                                                            ò
                                                                                                                                                                                        g
                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                       483
                                                                                                                                                                                                                                                                                                                                                                                   543
                                                                                                                                                                                                                                                                                             363
                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                          423
                                                                                                                                                                                                                                                                                                                                                                                                  243
                                                                                                                                                                                                                                                                303
                                                                                                                                                                                                                                                                                101
                                                                                                                                                                                                                                                                                                                                          ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                     4.1
                                                                                                                                                                                                                       AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                                  GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                          22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAT
       1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
Patent: US 6211338-A 99 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                               ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                      651
168
17
7
3
                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 bp
6211338.
                                                                                                                                  Gaps:
                                                                                                                                                (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          651
Sequence 93 from patent US 62
AR145252
AR145252.1 GI:15107119
                                           Location/Qualifiers
1. .651
                                                          /organism-"unknown"
187 c 200 g
                                                                                                                                               US-09-965-594-16 (1-197) x AR145258
                                                                                             1.83e-66
901.50
94.878
86.158
87.358
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (bases 1 to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unclassified.
Unclassified
                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                                  120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown
                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                              2
                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                                                                            102
                                                                                                                                                                                                                                                                                                                          364
                                                                                                                                                                                                                                                                                                                                          122
                                                                                                                                                                                                                                                                                                                                                       424
                                                                                                                                                                                                                                                                                                                                                                      142
                                                                                                                                                                                                                                                                                                                                                                                   484
                                                                                                                                                                                                                                                                                                                                                                                                  162
                                                                                                                                                                                                                                                                                                                                                                                                                               182
                                                                                                                                                                                                                                                                                                                                                                                                                                             604
                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
AR145252
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                    source
                                                                 BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                      JOURNAL
                AUTHORS
                      TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                            a
                                                                                                                                                                                                      g
                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                              ò
                                                                                                                                                                                           õ
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                  423
                                                                                                                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                                                                                                                                                                                                                                                                                                                        483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                543
                                                                                                                                                                                                                                                                                               243
                                                                                                                                                                                                                                                                                                                                      303
                                                                                                                                                                                                                                                                                                                                                          101
                                                                                                                                                                                                                                                                                                                                                                            363
                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                                                           162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                        4
                                                                                                                                                                                                                                                  ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      544 TCGGGGCACGCTGTGGGCATCTTCCGGCCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                                                                                                                                                                                                        22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                 GlySerValValIleValGlyArglleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 (Dases 1 to 651)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
 Yao, N.
is C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, Single-Chain recombinant complexes of hepatitis C protease and NS4A cofactor peptide Patent: US 621138-A 93 03-APR-2001; Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   651
169
15
8
8
                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                      145
                                                                                                                                                            Gaps:
                                                                                                                                                                               US-09-965-594-16 (1-197) x AR145252 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 94 from patent US 62.
AR145253
AR145253.1 GI:15107120
                                                          /organism="unknown"
188 c 199 g
                                                                                                         2.21e-66
900.50
94.36%
86.67%
87.26%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unclassified.
                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown.
                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                          82
                                                                                                                                                                                                                                                                                                                                                                           304
                                                                                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                                                                                                                                   364
                                                                                                                                                                                                                                                                                                                                                                                                                                       122
                                                                                                                                                                                                                                                                                                                                                                                                                                                         424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             604
                                                                                                            No.:
                                                                   BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
TITLE
                             JOURNAL
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 14
AR145253
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
  AUTHORS
          TITLE
                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                      ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
```

σ

```
completed: August 31, 2003, 00:46:14
ne : 2570.57 secs
                                                                 Percent Similarity:
Best Local Similarity:
                       ಡ
                                             Scores:
                                                                                                                                                                                                                                                                                              122
                                                                                 Query Match:
         source
                                                   Pred. No.:
                       COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                        Search con
Job time
                                            Alignment
  PEATURES
                              ORIGIN
                       BASE
                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                   ò
                                                                                                                                q
                                                                                                                                                ò
                                                                                                                                                           qq
                                                                                                                                                                            ò
                                                                                                                                                                                           q
                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                               PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                           243
                                                                                                                                                                                                                      ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                               41
                                                                                                                                                                                           61
                                                                                                                                                               22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                           1 (bases 1 to 1998)
Malcolm.B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 106 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                 linear
                                                                 651
169
15
8
                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145265 1998 bp C
Sequence 106 from patent US 62113338
AR145265 GI:15107132
                                                                 Length:
Matches:
protease and NS4A cofactor peptide
Patent: US 6211338-A 94 03-APR-2001
Location/Qualifiers
                                      145
                                                                                                      Gaps:
                                                                                                                    JS-09-965-594-16 (1-197) x AR145253 (1-651)
                             /organism="unknown"
188 c 199 g
                                                                2.21e-66
900.50
94.36%
86.67%
87.26%
                       1. .651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unclassified.
                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown
                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                    544
                                                                                                                                                64
                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                      244
                                                                                                                                                                                                                                                                  304
                                                                                                                                                                                                                                                                                 102
                                                                                                                                                                                                                                                                                              364
                                                                                                                                                                                                                                                                                                             122
                                                                                                                                                                                                                                                                                                                           424
                                                                                                                                                                                                                                                                                                                                          142
                                                                                                                                                                                                                                                                                                                                                       484
                                                                                                                                                                                                                                                                                                                                                                      162
                                                                                                                                                                                                                                                                                                                                                                                                   182
                                                                                                                                                                                                                                                                                                                                                                                                                 604
                                                                                                                                                                                                                                                   82
                                                                                               Query Match:
OB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                    DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                         source
                                                                  Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
TITLE
         JOURNAL
                                      SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                        AR145265
                                             DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                               LOCUS
                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                   Š
                                                                                                                                             g
                                                                                                                                                          5
                                                                                                                                                                        Q
                                                                                                                                                                                           රු යු
                                                                                                                                                                                                                  2 8 2 8
                                                                                                                                                                                                                                                                            à d
                                                                                                                                                                                                                                                                                                             50
                                                                                                                                                                                                                                                                                                                                          ò
```

```
363
                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                                                                                                                                                                                                                                                                                                                            483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                603
                                                                                                                                                                         183
                                                                                                                                                                                                                                                                                    243
                                                                                                                                                                                                                                                                                                                             303
                                                                                                                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                                                                                                                                                  364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                     41
                                                                                                                                                                                                                              124 TCCCAACAGACGCGGGGCCTACTTGCTTGCATCATCACTAGCCTTACAGGCCGGGACAAG
                                                                                                                                                                                                                                                                          42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                          62 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                     ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                     22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                               u
                               425
                                                                                                                               Gaps:
                                                                                                                                                     US-09-965-594-16 (1-197) x AR145265 (1-1998)
Location/Qualifiers
                   /organism="unknown"
597 c 567 g
                                                                         7.61e-66
900.50
93.88%
86.22%
           . 1998
```

Run on:

and is derived by analysis of the total score distribution

```
Hepatitis C virus Annal Complex NS4A-NS3 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDNA encoding hepa
Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Combined open read
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Combined open read
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Nucleotide sequenc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Composite hepatiti
                                                                                                                                                                                                                                   Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALIGNMENTS
                                              SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80363
AAX80347
AAX80348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAX80361
AAX80365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAN90336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                aax80357
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA73331 standard; DNA; 594
                                                                                                                                                              В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000 (first entry)
                                                                                 %
Query
Match Length I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               612
651
1998
1998
1998
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       651
651
1998
2016
648
650
650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus. Synthetic.
                                                                                                                                                                                                                                100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 903.5
901.5
900.5
900.5
900.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               904.5
904.5
904.5
904.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  897.5
897.5
897.5
896.5
894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                893.5
893.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8888.5
8888.5
8888.5
888.5
884.5
                                                                                                                                                              Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA73331
IID AAA7
XX
XX
AC AAA7
XX
XX
DT 19-D
DE HEPA
XX
KW HEPA
KW HEPA
KW IIVE
XX
KW I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Command line parameters:
-MODEL-framet-p2n.model -DEV-x1p
-MODEL-framet-p2n.model -DEV-x1p
-MODEL-framet-p2n.model -DEV-x1p
-WO-Cgn2_1/USPTO_spool/US09955594/runat_29082003_151918_28302/app_query.fasta_1.2872
-DB-N_Geneseq_19Jun03 -OFHT-fastap -SUFFIX-np -MINMATCH-0.1 -LOOPCL-0
-LOOPEXT-0 -UNITS-bits -STAFT-1 -FMATRIX-blosum62 -TRANS-human40.cdi
-LIST-45 -DCALIGN-200 -THR_SCORE-pct -THR_MAX_100 -THR_MIN-0 -ALIGN-15
-MODE-LOCAL -OUTPHT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0 -MAXLEN-200000000
-USER-DS0995554_eCGN_1_1142_ernat_29082003_151918_28302 -NCPU-6 -ICPU-3
-NO_MMAP -LARGEQUERY -NEG_SCORE-0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_ITMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT=0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT=0.5 -DELDP-6 -DELEXT-7
                                                                                                                                                                                                                                                              (without alignments)
2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             / S1DS1/gcgdata/geneseq/penesegn-embl/Mal981.DAT:
/ S1DS1/gcgdata/geneseqy.enesegn-embl/Mal982.DAT:
/ S1DS1/gcgdata/geneseqy.embl/Mal983.DAT:
/ S1DS1/gcgdata/geneseqy.enesegn-embl/Mal985.DAT:
/ S1DS1/gcgdata/geneseqy.enesegn-embl/Mal985.DAT:
/ S1DS1/gcgdata/geneseqy.enesegn-embl/Mal985.DAT:
/ S1DS1/gcgdata/geneseqy.embl/Mal986.DAT:
/ S1DS1/gcgdata/geneseqy.enesegn-embl/Mal989.DAT:
/ S1DS1/gcgdata/geneseqy.enesegn-embl/Mal989.DAT:
/ S1DS1/gcgdata/geneseqy.eneseqn-embl/Mal990.DAT:
/ S1DS1/gcgdata/geneseqy.enespn-embl/Mal990.DAT:
/ S1DS1/gcgdata/geneseqy.eneseqn-embl/Mal991.DAT:
/ S1DS1/gcgdata/geneseqy.eneseqn-embl/Mal991.DAT:
/ S1DS1/gcgdata/geneseqy.eneseqn-embl/Mal993.DAT:
/ S1DS1/gcgdata/geneseqy.eneseqn-embl/Mal993.DAT:
/ S1DS1/gcgdata/geneseqy.eneseqn-embl/Mal993.DAT:
/ S1DS1/gcgdata/geneseqy.eneseqn-embl/Mal999.DAT:
/ S1DS1/gcgdata/geneseqy/geneseqn-embl/Mal999.DAT:
/ S1DS1/gcgdata/geneseqy/geneseqn-embl/Mal909.DAT:
/ S1DS1/gcgdata/geneseqy/geneseqn-embl/Mal909.DAT:
/ S1DS1/gcgdata/geneseqy/geneseqn-embl/Mal999.DAT:

                                                                                                                                                                                                                                                                                                                                                                                                                                            ......VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                    August 30, 2003, 19:13:57; Search time 182.939 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /SIDS1/gcgdata/geneseg/genesegn-embl/NA2002.DAT:*/SIDS1/gcgdata/geneseg/genesegn-embl/NA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /SIDŠĪ/gcgdata/geneseq/geneseqn-embl/Na1980.DAT:
/SIDS1/gcgdata/geneseq/geneseqn-embl/Na1981.DAT:
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

    nucleic search, using frame_plus_p2n model

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2552756 seqs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  summaries
                                                                                                                                                                                                                                                                                                                                                                                                         1032
1 MKKKGSVVIVGRINLSGDTA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Xgapext
Ygapext
Fgapext
Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                       US-09-965-554-16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BLOSUM62
Xgapop 10.0 , X
Ygapop 10.0 , Y
Fgapop 6.0 , F
Delop 6.0 , D
                                                                                                                                                                                                                                                                                                                                                                       Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Searched:
```

us-09-965-594-16.rng

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                   "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                                                 Goldfarb V;
                                                                                                                                                                                                                                                   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                   Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                   Wittekind M, Weinheimer S,
                                                                                                                                             2000WO-US00345
                                                                                                                                                                               99US-0115271
                    /*tag= a
/product= '
1..594
/*tag=
                                                                                                                                                                                                                                                                                     WPI; 2000-465976/40.
P-PSDB; AAB15222.
                                                                       WO200040707-A1
                                                                                                                                                                               08-JAN-1999;
                                                                                                                                             36-JAN-2000;
```

Sequence 594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;

594 197 0 0 0 Conservative: Mismatches: Indels: Length: Matches: Gaps: 8.48e-88 1032.00 100.00% 100.00% 100.00% Percent Similarity: Best Local Similarity: Alignment Scores:

US-09-965-594-16 (1-197) x AAA73331 (1-594) ò

Q ò g

0, TyralaglnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp 61

9

LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla 60 41

a

ç

셤 ò å

ð

180

- ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80 61
- 81
- GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120 101

301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC 360 480 180 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix ${\bf 0}$ amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis ${\bf C}$ LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 481 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds. AlavalAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197 591 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3. /product= "NS4A-NS3 fusion protein #3" Wittekind M, Weinhelmer S, Zhang Y, Goldfarb V; Location/Qualifiers 1..594 /*tag= a 8 BP (BRIM) BRISTOL-MYERS SQUIBB 99US-0115271 06-JAN-2000; 2000WO-US00345 AAA73330 standard; DNA; 594 WPI; 2000-465976/40. P-PSDB; AAB15221. Hepatitis C virus. WO200040707-A1 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Synthetic 141 161 AAA73330; 181 121 RESULT 2 AAA73330 g ð ò g õ õ à

Claim 26; Fig 13; 66pp; English.

The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.

ò

ద ò qq දු

ò

a õ g õ

ò

m

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver ancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 LeuvalThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 MetLysLysLysClySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
                                                                                                                                    Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-16 (1-197) x AAA73332 (1-594)
                                                                                                                                    2hang Y,
                                                                                                             ဗွ
                                                                                                                                                                                                                                                             Claim 26; Fig 15; 66pp; English
                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5.51e-85
1002.00
98.48%
98.48%
97.09%
                                                                                                                                    Weinheimer S,
                                                           06-JAN-2000; 2000WO-DS00345
                                                                                   99US-0115271
                                                                                                                                                          WPI; 2000-465976/40.
P-PSDB; AAB15223.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
          WO200040707-A1
                                                                                                                                    Wittekind M,
                                                                                   08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                 13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ..
₽
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         540
                                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                                           100
                                                                                                                                                                                                                                                                                                                                                                                    300
                                                                                                                                                                                                                                                                                                                                                                                                             120
                                                                                                                                                                                                                                                                                                                                                                                                                       ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                        180
                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                           80
                                                                                                                                                                                                            40
                                                                                                                                                          ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                            TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                           LysasnGlnValGluGlyGluValGlnIleValSerThralaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                         GlaAlaProGlaGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /product= "NS4A-NS3 fusion protein #5"
          other;
                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
          0
          Ë
         G; 149
                                                                                                            Gaps:
                                                                                                                                    US-09-965-594-16 (1-197) x AAA73330 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
          C; 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ВP
                                            3.33e-86
1015.00
98.48%
98.48%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA; 594
          A; 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ø
         103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA73332 standard;
                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
         BP;
         Sequence 594
                                 Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic
                                                                                                                                                                                                                                                           41
                                                                                                                                                                                                            21
                                                                                                                                                                                                                                                                                    7
                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                    241
                                                                                                                                                                                                                                                                                                                                                                                                           101
                                                                                                                                                                                                                                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     541
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA73332;
                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAA73332
                                                             Score:
```

g

ö 셤 à a ò 120

20 9 9

셤

셤 à 9

180

240 100 300

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A processe enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Ntver cancer. The present invention concerns a number of NS3 inhibitors of this type, as well as enabling structural studies of the protease and protease einhabitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                         LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                        ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
421 CIGCIGICCCCCCCCCCAICTCCIACCIGAAAGGIICCICCGGIGGICCGCIGCIGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag= a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                     161
                                                                                                                                                                                                                                                           481
                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA73333;
                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SON X HE FERRY SON X REAL FERRY SON X RE
                                                                                                                                                                                                                                                           g
                             셤
                                                                                      ò
                                                                                                                                                                                                                                                                                                                   ö
```

Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0 other;

```
240
                                                                                                                                                                                                                                                                                                                                120
                                                                                                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                                                                                                            140
                                                                                                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                                                                       160
                                                                                                                                                         180
                                                                                                                                                                                                                                                                                     100
                                                                                                                                                                                                                                                                                                          300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                0
                                                                                                                                                                                             9
                                                                                                                                                                                                                                        80
                                                                                                     20
                                                                                                                          9
                                                                                                                                                                                         1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                              21 TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                    SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                 301 CAGGCICCGCAGGGIICCCGIICCCIGACCCCGIGCACCIGCGGIICCICCGACCIGIAC
                                                                                                                                                                                                                                                                                                                                                                                                                      LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                   481 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTTCCACCCGTGGTGTTGCTAAA
                                                                                                               ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                          LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /*tag= a
/product= "NS4A-NS3 fusion protein #2"
 594
191
2
0
0
                      Conservative:
Mismatches:
Indels:
  Length:
Matches:
                                                         Gaps:
                                                                             US-09-965-594-16 (1-197) x AAA73333 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ВР
7.34e-84
990.00
97.97%
96.95%
95.93%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
Synthetic.
                                 Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200040707-A1
                       Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000
                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                    83
                                                                                                                                                                                                                                                                                                                               101
                                                                                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAA73329
                                              Query Match:
                                  Best Local
 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA73329
IID AAA7
IID AAA7
XX AC AAA7
XXX
XXX
XXX
XXX
XXX
XXX
XXX
HEPP
XXX
HEPP
XXX
HEPP
XXX
HEPP
XXX
HEPP
T
COS
Syntt
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Key
```

S

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the Polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 martants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                             Claim 26; Fig 12; 66pp; English
                                                                                     (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                  Wittekind M, Weinheimer S,
06-JAN-2000; 2000WO-US00345
                                             99US-0115271
                                                                                                                                                                             WPI; 2000-465976/40.
                                                                                                                                                                                                       P-PSDB; AAB15220
                                          08-JAN-1999;
```

Goldfarb V;

Zhang Y,

8

Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;

21 TyralaginglnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp 40 [[[[[[[[[[[]]]]]]]]]]]] 55 TACGCTCAGACCGACCGGTGGAGGGTTGCCAAGAAACCTCCAGACCGGTGGTGA 114 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80 1 MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla 20 1 ATGAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATACTACTGAACGAT-----GCT Length:
Matches:
Conservative:
Mismatches:
Indels: Gaps: (1-588)US-09-965-594-16 (1-197) x AAA73329 5.06e-83 981.00 96.95% 96.45% 95.06% Percent Similarity: Best Local Similarity: Query Match: Alignment Scores: 61 175 81 101 121 355 Pred. No.: Score: a g ð 셤 à a 6 g ò à č g à

The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enizymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant. Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C Hepatitis, NS3 protease, viral replication; chronic liver disease, liver failure, liver cancer; mutant; muteln; ds. Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7 "NS4A-NS3 fusion protein #7" Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other; 594 Goldfarb Zhang Y, Location/Qualiflers Claim 26; Fig 17; 66pp; English. 8 (BRIM) BRISTOL-MYERS SQUIBB Wittekind M, Weinheimer S, 6.36e-83 980.00 99US-0115271. 06-JAN-2000; 2000WO-US00345 AAA73334 standard; DNA; 594 /*tag= a /product= (first entry) 1..594 /*tag= WPI; 2000-465976/40. Hepatitis C virus. P-PSDB; AAB15225. WO200040707-A1 Alignment Scores: Pred. No.: Score: 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Synthetic AAA73334; RESULT a ò 셤 ò

```
240
                                                                                                                                                                                                                 120
                                                                                                                                                                                                                                                               140
                                                                                                                                                                                                                                                                                                   LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                    540
                                                                                                              180
                                                                                                                                                                                                                                                                                                                                                CTGCTGTCCCCGCGTCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGCTGTGC 480
                                                                                                                                                                                                                                                                                                                                                                  ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                  09
                                                                                                    40
                                                                                                                                         9
                                                                                                                                                                              80
                                                                 20
                                                                                                                                          LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                    ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                        GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                        LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                           MetLysLysClysClySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                          TyralacinginThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     itis; NS3 protease; viral replication; chronic liver disease;
failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                       AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /*tag= a
/product= "NS4A-NS3 fusion protein #8"
 00010
 Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ВР
                                                US-09-965-594-16 (1-197) x AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
97.46%
96.45%
94.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1..594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus.
Synthetic.
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                       61
                                                                                                                                                                              61
                                                                                                                                                                                                                                    241
                                                                                                                                                                                                                                                                        301
                                                                                                                                                                                                                                                                                                           361
                                                                                                                                                                                                                                                                                                                                                                                   481
                                                                                                                                          41
                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                              141
                                                                                                                                                                                                                                                                                                                                                421
                                                                                                                                                                                                                                                                                                                                                                  161
                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA73335;
                                                                                                                                                                                                                                                      101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                liver
                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                  ö
                                                                                   셤
                                                                                                     ŝ
                                                                                                                      a
                                                                                                                                         ð
                                                                                                                                                         g
                                                                                                                                                                                             đ
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                       ਨੇ
                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                              ç
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    301 CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80
                                                                                                                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 TyralaGlnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThr11eAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 ACCIGCATCAACGGIGITIGCIGGACCGITIACCACGGIGCIGGIACCGGIACCAICGCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   121 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
                                                                           Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-16 (1-197) x AAA73335 (1-594)
                                                                           Zhang Y,
                                                                                                                                                                                                                                     Disclosure; Fig 18; 66pp; English.
                                              ပ္ပ
                                           (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1.51e-82
976.00
95.94%
95.94%
                                                                           Weinheimer S,
             990S-0115271
                                                                                                         WPI; 2000-465976/40.
P-PSDB; AAB15226.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
Pred. No.:
             08-JAN-1999;
                                                                           Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
US-09-965-594-16 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            04-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                    415
                                                                                                                                                                                                                                                                                                                                                                                                          161
                                                                                                                                                                                                                                                                                                                                                                                                                                  475
                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              535
                                                                                                                                                                                                                                                                                                                                                             141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABA95615;
                                                                                                                                                                                                                   81
                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                          Š
                                              g
                                                                        ô
                                                                                            g
                                                                                                                       ò
                                                                                                                                             셤
                                                                                                                                                                     ò
                                                                                                                                                                                          g
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                         ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
ProAlaClyHisAlaValClyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                  197
                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                  AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                /*tag= a
/product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zhang Y,
                                                                                                                                                                                                                                                                                        Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                        BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protease:inhibitor complexes,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2.29e-79
942.00
94.42%
93.91%
91.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M, Weinheimer S,
                                                                                                                        588
                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                     (first entry)
                                                                                                                       AAA73328 standard; DNA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-465976/40.
P-PSDB; AAB15212.
                                                                                                                                                                                                                                                     Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000
  161
                         181
                                                  181
                                                                        541
                                                                                                                                              AAA73328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
Q
                                                                                               RESULT 8
AAA73328
                                                                                                                        g
                                                   ô
  6
                          a
```

Matches: Conservative: Mismatches: Indels:

Percent Similarity: Best Local Similarity: Query Match: DB:

Score:

```
ProalaglyHisalaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                             534
                                                                                SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                 LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                          20
                                                                                                                           80
                    54
                                         40
                                                                                                                                                                                                                                                                                                                                                     CCGCTGGTCACCTGTTGTTTTTTCCGTGCTGCTGTTTGCACCCGTGGTGCTTAAA
                                                                                                                           ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                        TyralaGlnGlnThrArgGlyGluGluGlyGySglnGluThrSerGlnThrGlyArgAsp
                    1 ATGAAAAAAAAGGTTCCGTTGTTATCGTCGGCCGTATAGTACTGAACGGT-----GCT
                                                                                                                                                                                                                                                                                                                                                                               Nucleic acid construct encoding chimeric Hepatitis C Virus (RCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ds,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pestivirus; Npro; protease; NS3; screening;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                BP
                                                                                                                                                                                                                                                                                                                                                                                                                                               ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  990S-0344456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99DS-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lai VCH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hong Z,
```

us-09-965-594-16.rng

```
pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
```

Example 2; Columns 17-28; 20pp; English.

The present invention relates to a nucleic acid construct encoding a colmaric Hepatitis C virus (HCV)-pestivitus genome. The construct comprises a pestivitus genome where a Npro pestivitus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease. Is replaced with a for screening compounds that Inhibit HCV in vivo by inhibiting HCV for screening compounds that Inhibit HCV in vivo by inhibiting HCV The present sequence is a chimaric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

413 GGTAGTGTTGTTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472 593 TGCATCAATGGGGTATGCTGGACTGTCTACCACGGGGCCGGAACGAGGACCATCGCATCA 652 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101 GlyServalvalIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr 21 22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 473 GCCCAGCAGACGAGAGGCCTCCTAGGGTGTAAGATCACCAGTCTGACTGGCCGGGACAAA 42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu CTTTCGCCCCGGCCCATTTCCTACCTAAAGGCTCCTCGGGGGGGTCCGCTGTTGTGCCCCC AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196 12734 180 4 8 3 Length:
Matches:
Conservative:
Mismatches:
Indels: US-09-965-594-16 (1-197) x ABA95615 (1-12734) Gaps: 7.1e-76 922.50 94.36% 92.31% Percent Similarity: Best Local Similarity: Alignment Scores: 182 'n 62 83 653 122 773 142 162 893 Query Match: Pred. No.: Score: 셤 ð a ò g g ô g 셤 à 셤 g ò ò à g õ ö

243

a

AAX80355 standard; cDNA; 1998

953

셤

RESULT 10 AAX80355 ID AAX8 XX

61

1998 170 16 7 3

```
64 GGTTCTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence encodes an example of the above complex. The present NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors by the protease activity and the APPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAG 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 AlaGlnGlnThrArgGlyGluGluGlvGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlySerValValIleValGlyArglleAsnLeuSerGlyAsp...--ThrAlaTyr
                                                                                        HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                          CDNA SEQ ID NO:105
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-16 (1-197) x AAX80355 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                     Weber PC,
                                                             complex encoding
                                                                                                                                                                                                                                                                                         98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.6e-76
911.50
94.90%
86.73%
88.32%
                                                                                                                                                                                                                                                              98WO-US24528
                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                     Taremi SS,
                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                               WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                           HCV NS4A-NS3
                                                                                                                                                                                                                                                                                         28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                WO9928482-A2
                                                                                                                                                                                                                                                            24-NOV-1998;
                                07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                     Malcolm BA,
                                                                                                                                                                                                                               10-JUN-1999
                                                                                                                                                                      Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62
AAX80355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
   832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
```

ż

σ

```
569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BP
 411 A; 595 C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80353 standard; cDNA; 1998
                            1.45e-75
908.50
94.90%
86.22%
88.03%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus Synthetic.
 BP;
                                                 Percent Similarity:
Best Local Similarity
Sequence 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999.
                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                    424
                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                                                                                                                                                                                              364
                                                                                                                                                                                                                                                                                                                                                                122
                                                                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                           484
                                                                                                                                                                                                                                                                                                                                                                                                                                                162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAX80353;
                                                                                                                      Ŋ
                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                      ö
                                                                                                                                        ద
                                                                                                                                                            õ
                                                                                                                                                                              음
                                                                                                                                                                                                    à
                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                 ô
                                                                            102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                          423
                                                                                                                                                            AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 protease domain. The present to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent covalent protease re more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                  HCV: hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                    ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                NS4A-NS3 complex encoding cDNA SEQ ID NO:109.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Weber PC,
                                                                                                                                                                                                                                                                AAX80359 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                        (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                   WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                       07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                               Synthetic
                                                                                                                                         484
                                                                                                                                                              162
                                                                                                                                                                                                      182
                                                                                                  424
                                                                                                                       142
                                                                                                                                                                                                                                                                                    AAX80359
 82
                     304
                                                           364
                                                                                                                                                                                                                                             RESULT 11
                                                                                                                                                                                                                                                                 a
                                        ò
                                                           g
                                                                                ŏ
                                                                                                  a
                                                                                                                       ö
                                                                                                                                          ద
                                                                                                                                                              ò
                                                                                                                                                                                 g
                                                                                                                                                                                                      õ
 õ
```

```
141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          483
                                                                                                                                                                                                                                                                                                              363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 243
                                                                                                                                                                                                                                                                                                                                                                       ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                   102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                      61
                                                                                                                                                                                          22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MCV: hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:103
                            Conservative:
Mismatches:
Indels:
Length:
Matches:
                                                                      Gaps:
                                                                                                    US-09-965-594-16 (1-197) x AAX80359 (1-1998)
```

us-09-965-594-16.rng

```
07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                   10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                     RESULT 13
                                                                                                                   AAX80354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 ValThrArgHisAlaAspValIleProValArgArgAlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                   The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent covalent protease remore soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlySerValValIleValGlyArglleAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1998
170
15
8
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                     ż
                                                                                                                                                                                          Disclosure; Page 160–162; 211pp; English.
                                                                                                                                                             New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (1-1998)
                                                                                                     Д,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-16 (1-197) x AAX80353
                                                                                                       Weber
                           98US-0094331
97US-0067315
98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                       1.8e-75
907.50
94.39%
86.73%
87.94%
                                                                                                     Taremi SS,
                                                                          (SCHE ) SCHERING CORP
                                                                                                                                 WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                    Malcolm BA,
                             28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       484
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
```

q ò a ò අ ò 셤 à

ģ a ö g õ g

ö

ó

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NWR spectroscopy. They can also be used for detecting inhibitors by the protease activity, the helicase activity and the APPase activity of NS3 The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 AlaGinGinThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
5 GlySerValVallleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease, hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                     NS4A-NS3 complexes are more soluble, stable and active the covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1998
170
15
8
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                              HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; Page 163-166; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-16 (1-197) x AAX80354 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ઇ
                                                                                                                                                                                                                               BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1998 BP; 410 A; 596 C;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weber
                                                                                                                                                                                                                            AAX80354 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          980S-0094331
970S-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1.8e-75
907.50
94.39%
86.73%
87.94%
                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
```

```
121
                                                                                                                                                                                                        ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                    ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                          AlaGlyHisAlaValGlyIlePheArgAlaAlaValCySThrArgGlyValAlaLySAla 181
                                   9
                                                                           CysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          has
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C; ds; viral prototoxophore; anti-viral; tumour;
virus; infection; antitumour; toxophore; human immunodeficiency virus;
HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                                                                                                              AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Novel synthetic viral prototoxophore for treating viral infections, toxin molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This invention relates to a novel synthetic viral prototoxophore
                                                                                                                                                                                                                                                                                                                                   ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                             Anti-viral synthetic prototoxophore associated DNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Shepard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 1; Page 62; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                   BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIC
                                                                                                                                                                                                                                                                                                                                                                                                  ABX15706 standard; DNA; 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26-APR-2002; 2002WO-US13223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27-APR-2001; 2001US-286893P
                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Cathers BE, Neuteboom
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WO200287500-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Jnidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                            28-MAR-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-NOV-2002,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      toxophore
                                                                                                                                                              102
                                                                                                                                                                                                        122
                                                                                                                                                                                                                          124
                                                                                                                                                                                                                                                                                         162
                                                                                                                                                                                                                                                                                                             544
                                                                                                                                                                                                                                                                                                                                   182
                                                                                                                                                                                                                                                                                                                                                      604
                                42
                                                                          62
                                                                                                                   82
                                                                                                                                         304
                                                                                                                                                                                  364
                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                    184
                                                                                                                                                                                                                                                                                                                                                                                                                        ABX15706;
                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                          a
                                                                                                                   ŏ
                                                                                                                                         g
                                                                                                                                                          Š
                                                                                                                                                                              g
                                                                                                                                                                                                      à
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                     ŏ
                                                    qq
                                                                          õ
                                ð
```

```
comprising a toxin molety operatively incorporated into a substrate domain specific for a viral enzyme. This prototoxophore may be bound and anodified by the viral enzyme thus converting it to a toxophore. Also disclosed in the invention is a method for enhancing the anti-viral effect of an antiviral agent, this method comprises contacting a cell, infected with a virus or is susceptible to infection, with a cell, of an invention further comprises an assay to identify an infected with a virus or is susceptible to infected cell with a candidate agent and comparing the ability of the agent to inhibit the candidate agent and comparing the ability of the agent to inhibit the crowth or infectivity of the virus in the cell. The prototoxophores of the invention may have virucide or antitumour activity. The prototoxophores of the invention may be useful for reducing or inhibiting viral infectivity, by contacting a cell (e.g. imphibory) neare cell, connective tissue cell, muscle cell or hepatocyte) which is infected with a virus or is susceptible to infection with a virus, with a capted to long term continuous culture or isolated from a subject. The prototoxophore is also useful for ameliorating the severity of a viral infection in a subject, where the virus (18vy, rhinovirus and immunodeficiency virus (11vy), herees simplex virus (18vy, rhinovirus and immunodeficiency virus (11vy), herees simplex virus (18vy, rhinovirus and antiviral prototoxophore associated DNA sequence represents an element of the subject. The present sequence represents a class as recombinant if a prototion are contactive and a recombinant if a class or serving and a recombinant is also every the present sequence is described as a recombinant prototoxophore is a recombinant in a prototion and a pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AlaGlyHisalaValGlyIlePheArgalaAlaValCysThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139 AACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCTGCCCAAACCTTCCTGGCAACG 198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        259 CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTAGACCAAGACCTTGTGGGCTGGCCC 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    the invention although it is clearly not a protein sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   612
178
3
11
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-16 (1-197) x ABX15706 (1-612)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7.89e-76
904.50
92.82%
91.28%
87.65%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

```
LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
124 ICCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                                                                            ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                          ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                            62 CyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                   ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                 82
                                                                                                                                                                                                                                          122
                                                                                                                                                                                                                                                                   424
                                                                                                                                                                                                                                                                                              142
                                                                                                                                                                                                                                                                                                                      484
                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                 g
                                                                                                                                                      g
                                                                                                                                                                                     ò
                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                              ò
                   ΟŸ
                                            g
                                                                         ò
                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 Perine protease domain, where the hydrophobic domain of native HCV NS3 Perine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more solubbe, stable and active than the noncovalent protease-peptide complexes previously available.
 499 GCGGGGCCACGCCGTGGGCCATATTTAGGGCCGCGGTGTGCACCCGTGGAGTGGCTAAGGCG 558
                                                                                                                                                                                                                HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                         Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                        HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Disclosure; Page 147-148; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ς,
                                                                                                           ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8.52c-76
904.50
94.87%
86.67%
87.65%
                                                                                                          AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                 98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                           98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                      97US-0067315
                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                 (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1999-385385/32
                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                         28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                   Synthetic.
                                                                                                                                     AAX80345;
                                                                                             AAX80345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred.
                                                                                                           a
 a
                             8
```

4 22

a ô

AlaGinGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys

GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr 21

US-09-965-594-16 (1-197) x AAX80345 (1-651)

Best Local Similarity:

Query Match:

Percent Similarity:

Search completed: August 30, 2003, 19:48:03 Job time : 188.939 secs

Title: Perfect score:

on:

Sequence:

Scoring table:

Searched:

19, Appl 19, Appl 19, Appl 2, Appl 1, Appli 2, Appli 2, Appli

13, Appl 16, Appl 1, Appli 8, Appli 15, Appli

Appl Appli Appli Appli Appli

Appl Appl Appl

```
Sequence 17, Application US/09965594
Fatent No. USZ0020106642A1
GENERAL INFORMATION:
APPLICANT: Witched, Michael
APPLICANT: Weinhelmer, Steven
APPLICANT: Coldiarb, Valentina
APPLICANT: Coldiarb, Valentina
APPLICANT: Goldfarb, Valentina
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: Facilitating Inhibitor Complexes
TITLE OF INVENTION: OF Protease:Inhibitor Complexes
TITLE OF INVENTION: OF Protease:Inhibitor Complexes
TITLE OF TREFFENCE: DB175equences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 1999-01-08
FRIOR FILING DATE: 1999-01-08
                                                                               Sequence 13, Appl
Sequence 13, Appl
Sequence 25, Appl
Sequence 4, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 1, Applequence 19, Sequence 19, Sequence 19, Sequence 17, Sequence 17, Sequence 17, Sequence 13, Sequence 13, Sequence 16, Sequence 16, Sequence 15, Sequence 15, Sequence 15, Sequence 69, Sequence 61, Apple 61, Sequence 61, Sequence 61, Sequence 61, Sequence 61, Sequence 61, Sequence 61, Apple 61, Sequence 61, Apple 61, Sequence 61, Apple 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2
Sequence 4
Sequence 6
Sequence 2
Sequence 2
Sequence 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Seguence
0. 0S-09-965-594-15
0. 0S-09-965-594-19
0. 0S-09-965-594-13
0. 0S-09-965-594-23
0. 0S-09-965-594-23
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
0. 0S-09-965-594-25
1. 0S-09-965-594-25
1. 0S-09-965-937-1
1. 0S-09-965-937-1
1. 0S-09-965-937-1
1. 0S-09-965-937-1
1. 0S-09-917-563-1
1. 0S-09-919-901-1
1. 0S-09-919-1
1. 0S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: DNA
CRGANISM: Hepatitis C virus
US-09-965-594-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                           612
9646
9646
9646
12980
12980
12980
9379
9416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        836
10803
10803
9416
9416
2061
13910
13910
2064
2523
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQ ID NO 17
LENGTH: 594
                                                                                                                                                                                                                                                                      881.5
881.5
881.5
878.5
878.5
878.8
878.8
878.8
863.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          863
863
860.5
           Command line parameters:
-MODEL-frame+_plan_model -DEV-xlp
-Q-Cqqq12_1/USFTO_spool/US09965594/runat_29082003_151920_28367/app_query.fasta_1.2872
-Q-/cqq12_1/USFTO_spool/US09965594/runat_29082003_151920_28367/app_query.fasta_1.2872
-Q-/cqq12_1/USFTO_spool/US09965594/runat_29082003_151920_28367
-UCOPCL-0 -LOOPEXT-0 -UNITS-bits -STARR-1 - BND-1 - MATRIX-blosum62
-TRANS-human40.cdl -LIST-45. DOCALIGN-200 -THR_SCORE-pet -THR_MAX.EN-CORE-500 -UNILEN-0
-MAXIEN-200000000 -USER-US09965594_GCGN_1_1_864_@runat_29082003_151920_28367
-NCPUG-6 -LCD-3 -NO_MAAP_LARREQUERY -REG_SCORE-0 - WALT -DSBLCCK-100
-LONGLOG -DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 - XGAPOP-10 -XGAPEXT-0.5
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 17, Appl
                                                                                                                                                                                   ; Search time 176.482 Seconds
(without alignments)
2560.981 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /cgn2_6/ptcdata/2/pubpna/US6U_NEW_PUB.seq:~
/cgn2_6/ptcdata/2/pubpna/US6U_PUBCOMB.seq:*
                              GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                    OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     594 10 US-09-965-594-17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1533700 seqs, 1147125425 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Published_Applications_NA:*
                                                                                                                                                                                      August 30, 2003, 19:26:03
                                                                                                                                                                                                                                                                                                                                                                                                                        Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                              US-09-965-594-16
1032
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query
Score Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                    BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1032 100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55....
1111....
115....
116....
```

Database :

ġ

8-09-965-594-15

Alignment Scores: Pred. No.: Score: Score: Sercent Similarity Best Local Similar Query Match: DB:	res: arity: milarity:	1.82e-110 1032.00 100.00% 100.00% 100.00%	Length: Matches: Conservative: Mismatches: Indels: Gaps:	594 197 0 0 0	
US-09-965-594	-16 (1-197	3-4-60-80 × (94-17 (1-594)		
0y 1 0b 1		LysGlySerValVal: 	I eValG yArg leA 	MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla 2	0 0
Oy 21		GlnThrArgGlyGlu(aluglyCysglnGlul 	TyralaginginThrargglygluglyglyglygluglytyselngluThrSerglnThrGlyargasp 4	20
Oy 41		ValGluGlyGluVal(SlnileValSerthra 	LysasnGlnvalGluGlyGluValGlnIleValSerThralaThrGlnThrPheLeuala 6 	0.08
Qy 61 Db 181		AsnGlyValCysTrp' - - - - - - - - - - - - - - - - - - -	ThrValTyrHisGlyA 	ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 8 	.0
Oy 81	2 — F	GlyProvalThrGlm 	MetTyrThrAsnVall 	erProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 	300
Oy 101 Db 301		GlnGlySerArgSer 	LeuThrProCysThro 	GlnalaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 1 	120 360
Oy 121 Db 361		Arghisalaaspval 	IeProvalArgArg 	LeuValthrarghisalaaspvalileProValargargargargglyaspserargGlyser 	140
Qy 141 Db 421		ProArgProlleSer 	TyrLeuLysglySer: 	LeuLeuSerProArgProlleSerTyrLeuLySg1ySerSerGlyGlyProLeuLeuCys	160
Oy 161 Db 481		HisalavalGlyIle 	PheArgAlaAlaVal(ProalaglyHisAlaValGlyIlePheArgalaAlaValCysThrArgGlyValAlaLys : 	180
Oy 181 Db 541		PhelleProValGlu TrCATCCCGGTTGAA	AlaValaspPhelleProValGluSerLeuGluThrThrMetArgSerPro 	MetArgSerPro 197 	
RESULT 2 US-09-965-59 Sequence 1 Sequence 1 Sequence 1 Sequence 1 Sequence 1 APPLICANT APPLICANT APPLICANT APPLICANT APPLICANT APPLICANT APPLICANT APPLICANT SEQUENCE SEE CURRENT AN CURREN	ULT 2 90-965-594-15 equence 15. Applicat: atent No. 10500201066 APPLICANT: Wittekind APPLICANT: Wittekind APPLICANT: Wittekind APPLICANT: Coldfarb TITLE OF INVENTION: WINDER OF INVENTION: CURRENT APPLICATION OF THE PRIOR FILING DATE: PRIOR APPLICATION NOWHER OF SEQ ID NOS SOFTWARE: PATENTION ON 15 LEMETH: 594 TYPE: DNA ORGANISM: HEPALLIS ORGANISM: HEPALLIS	ion US/09965594 642Al i. Michael ier, Steven agun Adultied Forms Facilitating I Sequences NUMER: US/09/9 1999-01-08 5: 26 fer. 2.0	RESULT 2 US-09-965-594-15 Sequence 15, Application US/09965594 Sequence 15, Application US/09965594 Parent No. US20020106642A1 GENERAL INFORMATION: APPLICANT: Wittekind, Michael APPLICANT: Weinheimer, Steven APPLICANT: Weinheimer, Steven APPLICANT: Goldfarb, Valentina APPLICANT: Goldfarb, Valentina TITLE OF INVENTION: Pacilitating Inhibitor Screening; TITLE OF INVENTION: Pacilitating Inhibitor Screening; TITLE OF INVENTION: Pacilitating Inhibitor Complexes FILE REFERENCE: DB17Sequences CURRENT APPLICATION NUMBER: 60/195/965,594 CURRENT FILING DATE: 1999-01-08 NUMBER: PATENTION NUMBER: 60/115,271 PRIOR FILING DATE: 1999-01-08 NUMBER: PATENTION NUMBER: 60/115,271 PRIOR FILING DATE: 1999-01-08 NUMBER: PATENTION OF PATENTION NUMBER: 60/115,271 FILENGTH: 594 TYPE: DATE APPLICATION CONTINES CORTWARE: PATENTIN CONTINES CORTWARE: PATENTIN CONTINES CORTWARE: PATENTIN CONTINES	C NS3 Protease for eening and Structural Studio lexes	ics

```
US-VG-9702-294-19
Sequence 19, Application US/09965594
Fatent No. US2020106642A1
Sequence 19, Application US/09965594
Fatent No. US2020106642A1
Sequence 19, Application:
APPLICANT: WitherLind, Michael
APPLICANT: Weinhelmer, Steven
APPLICANT: Goldfarb, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: OF Protease:Inhibitor Complexes
TITLE OF INVENTION: DBJ7Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT APPLICATION NUMBER: 60/115,271
FRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PATENTING DATE: 1999-01-08
SOFTWARE: PATENTING DATE: 1999-01-08
SOFTWARE: PATENTING DATE: 1999-01-08
LENGTH: 594
                                                                                                                                                                                                                              140
                                                                                                                                                                                                                                                                                                                                      180
                                                                                                                                                                                                                                                                                                                                                                                           540
                                                                                                                                                                   240
                                                                                                                                                                                                                                                                                                                      420
                                                                                                                      9
                                                                                                                                                                                           9
                                                                                                                                                  LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                    AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                     594
194
0
0
0
0
      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                    S-09-965-594-16 (1-197) x US-09-965-594-15 (1-594)
    1.7e-108
1015.00
98.48%
98.35%
                        ercent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-19
lignment Scores
                                                                                                                                                           41
                                                                                                                                                                                             61
                                                                                                                                                                                                             181
                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                 241
                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                      361
                                                                                                                                                                                                                                                                                                                                       141
                                                                                                                                                                                                                                                                                                                                                        421
                                                                                                                                                                                                                                                                                                                                                                                                             181
                                           uery Match:
         ..
92
                                                                                      ⋩
                                                                                                     a
                                                                                                                       ×.
                                                                                                                                      ð
                                                                                                                                                                      용
                                                                                                                                                                                          ₹ 8
                                                                                                                                                                                                                           č
                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                              ≿
                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                    2 8 2
                                                                                                                                                                                                                                                                                                                                                    8 3
                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                         Š
```

```
셤
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                          d
                                                                                                                                ò
                                                                                                                                            셤
                                                                                                                                                             ò
                                                                                                                                                                          셤
                                                                                                                                                                                          ò
                                                                                                                                                                                                           合
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                           dy
Qy
                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 21, Application US/09965594
| Sequence 21, Application US/09965594
| Patent No. US20020106642A1
| GENERAL INFORMATION:
| APPLICANT: Wittekind, Michael
| APPLICANT: Wittekind, Michael
| APPLICANT: The Inhelmer, Steven
| APPLICANT: Coldfarb, Walentina
| TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies
| TITLE OF INVENTION: Pacilitating Inhibitor Complexes
| TITLE OF INVENTION: Pacilitating Inhibitor Complexes
| TITLE OF INVENTION: US/09/965,594
| CURRENT APPLICATION NUMBER: US/09/965,594
| CURRENT APPLICATION NUMBER: 60/115,271
| PRIOR FILING DATE: 1999-01-08
| NUMBER OF SEQ ID NOS: 26
| NUMBER OF SEQ ID NOS: 26
| SOFTWAND FROM TOWN OF THE NOS: 26
                                                                                                                                                                                                                                                                                                                                           240
                                                                                                                                                                                                                                             180
                                                                                                                                                                                                                                                                                                                                                                                            540
                                                                                                                                                                                              180
                                                                                                                                                               TACGCTCAGCAGACTCGAGGTGAGGGTTGCCAAGAAACCTCCCAGACGGTCGTGAC 120
                                                                                                                                               TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp 40
                                                                                                                                                                               9
                                                                                                                                                                                                               80
                                                                                                                20
                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                            LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                       594
194
0
3
0
                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                US-09-965-594-16 (1-197) x US-09-965-594-19 (1-594)
                                                                                Gaps:
                                      5.45e-107
1002.00
98.48%
98.48%
97.09%
; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-965-594-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOFTWARE: Patentin Ver. 2.0
                                               Score:
Percent Similarity:
Best Local Similarity
                                 Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 4
US-09-965-594-21
                                                                                                                                                                                                                                                                                                                                                                                            481
                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                               61
                                                                                                                                                                                                                                                             241
                                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                21
                                                                                                                                                                               41
                                                                                                                                                                                              121
                                                                                                                                                                                                               61
                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                              101
                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                             421
                                                                        Query Match:
DB:
                                                                                                                õ
                                                                                                                               g
                                                                                                                                                ô
                                                                                                                                                               g
                                                                                                                                                                                           a
                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                        ô
                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                           οy
                                                                                                                                                                                                         ò
                                                                                                                                                                                                                          a
```

```
Studies
                                                                                                                                                                                                                                                                                                                                                                                                                          240
                                                                                                                                                                                                                                                                                                                                80
                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 13, Application US/09965594

Patent No. US20020106642a1

GENERAL INFORMATION:
APPLICANT: Withhelmer, Steven
APPLICANT: Weinheimer, Steven
APPLICANT: Weinheimer, Steven
APPLICANT: Weinheimer, Steven
APPLICANT: Chang, Yaginn
APPLICANT: Chang, Yaginn
APPLICANT: Chang, Yaginn
TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Stud
TITLE OF INVENTION: OF Protease:Inhibitor Complexes
TITLE OF INVENTION: OF Protease:Inhibitor Complexes
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT APPLICATION NUMBER: 60/115,271
PRIOR PILING DATE: 1999-01-08
                                                                                                                                                                                                                                                                                                                                             MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                     61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              594
191
0
0
                                                                                              Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                  US-09-965-594-16 (1-197) x US-09-965-594-21 (1-594)
                                                                                    Length:
                                                                                  1.34e-105
990.00
97.97%
96.95%
95.93%
                                      virus
; SEQ ID NO 21
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Hepatitis C
US-09-965-594-21
                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                     Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-13
                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                                                                                                                                   241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        141
                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            421
```

```
) ORGANISM: Hepatitis
US-09-965-594-23
                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 7
US-09-965-594-25
                                                                        Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                              301
                                                                                                                                                                                                                                                                                                                                                                          421
                                                                                                                                                                              21
                                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                                                                  241
                                                                                                                                                                                                                                                                                                101
                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                            361
                                                                                                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                                       481
                                                                                       Score:
                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                      ć
                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                            â
                                                                                                                                                                           ŏ
                                                                                                                                                                                         셤
                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                               요
                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                 ò
                                                                                                                                                                                                     ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 23, Application US/09965594

Sequence 23, Application US/09965594

GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Witchell
APPLICANT: Cladge Yaqun
APPLICANT: Coldach, Valentina
APPLICANT: Coldach, Valentina
TITLE OF INVENTION: Modified Forms of Repatitis C NS3 Protease for TITLE OF INVENTION: Modified Forms of Repatiting Inhibitor Screening and Structural Studies; TITLE OF INVENTION: of Protease:Inhibitor Complexes
FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
                                                                                                                                                                                                                                                                                                                140
                                                                                                                                                                                                                                                                                                                                             160
                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                              SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                    294
                                                                                                                                                                                                                                                                                  414
                                                                                                                                                                                                                                                                                                                                                           474
                                                                                                                                                                                                                                                                                                                                                                                         534
                                                                                                                                                                                                                                        234
                                                                                                                                                                                                          174
                                                                                                                                                                                                                         80
                                                                                                                                   50
                                                                                                                                                                                            9
                                                                                                                                                54
                                                                                                                                                                                          LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                      588
1190
12
                                                                                Conservative:
Mismatches:
Indels:
                                                                                                                    US-09-965-594-16 (1-197) x US-09-965-594-13 (1-588)
                                                                 Length:
Matches:
                                                                                                      Gaps:
                                                                 1.46e-104
981.00
96.95%
96.45%
95.06%
                                     C virus
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver.
SEQ ID NO 13
LENGTH: 588
                                                                                                      10
                           ; TYPE: DNA
; ORGANISM: Hepatitis
US-09-965-594-13
                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 6
US-09-965-594-23
                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                           191
                                                                                                                                                                                             41
                                                                                                                                                                                                           115
                                                                                                                                                                                                                          6.1
                                                                                                                                                                                                                                        175
                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                    235
                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                              355
                                                                                                                                                                                                                                                                                                                                             141
                                                                                                                                                                                                                                                                                                                                                           415
                                                                                                                                                                                                                                                                                                                                                                                         475
                                                                          Score:
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                   ò
                                                                                                                                                g
                                                                                                                                                                                            δ
                                                                                                                                                                                                        q
                                                                                                                                                                ò
                                                                                                                                                                           Q
```

```
Protease for and Structural Studies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerProLysGlyProvalThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuLeuSerProArgPro11eSerTyrLeuLysG1ySerSerG1yG1yProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGCTGTTTCCACCCGTGGTGTTGCTAAA
                                                                                                                                                                                                                                                                                                                      CTGCTGTCCCCCGCGTCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGCTGTGC
                                                                                                                                                                                                                                                                                                                                                                                     TyralaGlnGlnThrargGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 25, Application US/09965594

Patent No. US20020106642A1

GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Withkind, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Weinheimer, Steven
APPLICANT: Goldfarb, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Prot
TITLE OF INVENTION: Facilitating Inhibitor Screening and
TITLE OF INVENTION: of Protease:Inhibitor Complexes
FILE REFERENCE: DB17Sequences
                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                         US-09-965-594-16 (1-197) x US-09-965-594-23 (1-594)
                                                                                                                                                                                                                                          Indels:
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PATENTIN Ver. 2.0
SEQ ID NO 23
LENGTH: 594
TYPE: DNA
                                                                                                                                                                            1.93e-104
980.00
97.46%
96.45%
                                                                                                                virus
                                                                                                                U
```

9

9 40 240 100 300

80

540

140

```
355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         535
                                                                                                                                                                                                                                                                          21
                                                                                                                                                                                      Ouery Match:
DB:
                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Zhang, Yaqun
APPLICANT: Zhang, Yaqun
APPLICANT: Goldfarb, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 540
                                                                                                                                                                                                                                                       240
                                                                                                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ProalaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                        SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                   80
                                                                                                                                                                                                                   20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 ATCADADADADADADAGATCCGTTGTTATCGTCGCCGTATCAACCTGTCCGGTGACCCCCT
                                                                                                                                                                                                                                                                                           LysasnClnValGluGlyGluValGln11eValSerThralaThrGlnThrPheLeuAla
                                                                                                                                                                                                                  MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                AAAAACCAGGTTGAAGGTGAAGTTCAGATCGTTTCCACCGCTGCTCAGACCTTCCTGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                            594
189
0
0
0
                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                US-09-965-594-16 (1-197) x US-09-965-594-25 (1-594)
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION UNDER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARR: PATENTIN Ver. 2.0
SOFTWARR: PATENTIN Ver. 2.0
SOFTWARR: PATENTIN Ver. 2.0
TENGTH: 594
TREE: DNA
TYPE: DNA
UNGANISM: Hepatitis C virus
US-09-965-594-25
                                                                                                                                                                   Indels:
                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 4, Application US/09965594 Patent No. US20020106642Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Wittekind, Michael APPLICANT: Weinheimer, Steven
                                                                                                                             5.62e-104
976.00
95.94%
95.94%
94.57%
                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
US-09-965-594-4
                                                                                                                                                                                                                                                         21
                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                             41
                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                           241
                                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           421
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                               Pred. No.:
                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                   ò
```

```
161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCysThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55 TACGCTCAGCAGACTCGAGGTCTGCTGGGTTGCATCATCACCTCCCTGACCGGTCGTGAC 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGCTGTTTGCACCCGTGGTGTTGCTAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                    TyralaGloGloThrArgGlyGluGluGlyCysGloGluThrSerGloThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: CATHERS, Brian
APPLICANT: NEUTEBOON, SASKia
APPLICANT: SHERRE, Nichael
ITILE OF INVENTION: VIRAL ENZYME ACTIVATED PROTOTOXOPHORES
                                                                                                                                                                                                                                   1185
129
129
 Complexes
                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                            US-09-965-594-16 (1-197) x US-09-965-594-4 (1-588)
Protease: Inhibitor
                                                                                                                                                                                                                                                                                                             Gaps:
           FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,59-
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 4
LENGTH: 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 6, Application US/10133133A Publication No. US20030114385A1 GENERAL INFORMATION:
                                                                                                                                                                                                                                 4.84e-100
942.00
94.42%
93.91%
                                                                                                                                                   TYPE: DNA ORGANISM: Hepatitis C virus
                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-10-133-133A-6
                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                     US-09-965-594-4
```

414

534

100

294

234

80

174 9

0.4

φ

```
RESULT 11
US-09-238-076-1
                                                                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                  142 LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               439 CTGTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCCTCGGGGGGTCCGCTGTTGTGCCCC 498
                                                                                                                                                                                                                                                                                                                                                                                                                   82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                              GlySerValVallleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                     22 AlaGinGinThrArgGlyGluGluGlyGycysGlbGluThrSerGinThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                            CysileAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
   VIRAL INFECTIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
TITLE OF INVENTION: AND USE OF SAME TO TREAT VIR FILE REFERENCE: NB 2021.00
CURRENT APPLICATION NUMBER: US/10/133,133A
CURRENT FILING DATE: 2002-04-26
PRIOR APPLICATION NUMBER: 60/286,983
PRIOR FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 6
                                                                                                                                                                                                                                US-09-965-594-16 (1-197) x US-10-133-133A-6 (1-612)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             02-09-742-659-3
; Sequence 3, Application US/09742659
; Patent No. US20010034019Al
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hong, Zhi
Butkiewicz, Nancy J.
Zhong, Weidong
Ingravallo, Paul
                                                                                                                                                         1.13e-95
904.50
92.82%
91.28%
                                                                                                                 ORGANISM: Hepatitus C. Virus
                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                 Alignment Scores:
Pred. No.:
                                                                                                                            US-10-133-133A-6
                                                                                              LENGTH: 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                         TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                            62
                                                                                                                                                                                                 Query Match:
                                                                                                                                                                    Score:
                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                      염
                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                               윱
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
```

```
3354 CGTAGGGCCAGGAGATACTGCTTGGACCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1834 TCGGGGGGTCCGCTGTTGTGCCCCGCGGGACACGCCGTGGGCCTATTCAGGGCCGCGGGTG 1893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3474 ACCAGTCTGACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATCGTGTCAACT 3533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34 ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                       9646
172
10
13
9
                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-16 (1-197) x US-09-742-659-3 (1-9646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 LysLysGlySerValVallleValGlyArglleAsn-
                                                     ON: Chimeric HCV/GBV-B viruses
ID01116
                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                        CURRENT APPLICATION NUMBER: US/09/742,659
CURRENT FILING DATE: 2000-12-21
PRIOR APPLICATION NUMBER: US 60/171,469
PRIOR FILING DATE: 1999-12-22
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
APPLICANT: Wright-Minoque, Jacquelyn
APPLICANT: Lau, Johnson Y.
APPLICANT: Lemon, Stanley M.
TITLE OF INVENTION: Chimeric HCV/GBV-
FILE REPERENCE: ID01116
                                                                                                                                                                                                                                                                                                                       2.78e-92
888.50
89.22*
84.31*
86.09*
                                                                                                                                                                                                                                                        virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       194 MetArgSerPro 197
                                                                                                                                                                                                                                      ; TYPE: DNA
; ORGANISM: Hepatitis
US-09-742-659-3
                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                     9646
                                                                                                                                                                                                                                                                                                                        Pred. No.:
```

```
AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                             CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/995,937
FILING DATE: 28-No. US20030028010A1-2001
CLASSIFICATION: <URknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,756
FILING DATE: 04-May-1998
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       et al.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGIH: 9646 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1, Application US/09995937; Publication No. US20030028010A1
GENERAL INFORMATION: APPLICANT: RICE, CHARLES et
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEPHONE: 314-727-5188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: 314-727-6092 INFORMATION FOR SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                      3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY: linear MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                                                                                                                                   194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-995-937-1
                   94
                                                                                               114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                             a
                                                                                             à
                                                                                                                             q
                                                                                                                                                                    õ
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3594 GCCGGAACGAGGACCATCGCATCACCCAAGGGTCCTGTCATCCAGATGTATACCAATGTG 3653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3414 TIGCIGGCGCCAICACGGCGTACGCCCAGCAGACGAGAGGCCICCTAGGGTGIATAAIC 3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal 93
                                                   APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS:
TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
MUMBER OF SEQUENCES: 21
CORRESPONDEMES: 21
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9646
172
10
13
9
                                                                                                                                                                                                                                          ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 LysLysGlySerValValIleValGlyArgIleAsn--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-16 (1-197) x US-09-238-076-1 (1-9646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ATTORNEY AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
FILING DATE:
Sequence 1, Application US/09238076
Patent No. US20020102540A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         888.50
89.22%
84.31%
86.09%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ANTI-SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-238-076-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
```

ô 요 ð 셤 ö g δ a ç

```
2.78e-92
888.50
89.22%
84.31%
                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                               US-09-917-563-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3534
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                           15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           154
                                                                                                                                                                                                                                                                                                                                                         Query Match:
DB:
                                                                                                                                                                                                                                                                                                                Pred. No.:
                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                    3833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3773
                                                                                                                                                                                                                                               354 GACCAAGACCTTGTGGGCTGGCCCGCTCCTCAAGGTTCCCGCTCATTGACACCCTGCACC 3713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                         3474 ACCAGCCTGACTGGCCGGGACAAAACCAAGTGGAGGGTGAGGTCCAGATCGTGTCAACT 3533
                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                         AspLysAspLeuValGlyTrpGlnalaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                                                                                                                                                                                                                                                                                                                               134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 153
                                                                                                                                                                                                                                                                                                                                                                     CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                              --- LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                                                                                                                                                   54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                ThrSerGlnThrGlyArgAspLySAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF
9646
172
10
13
 Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
ADDRESSE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
                                                                                                    LysLysGlySerValVallleValGlyArglleAsn-
                                                                             US-09-965-594-16 (1-197) x US-09-995-937-1 (1-9646)
                                  Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   et al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 1, Application US/09917563
Publication No. US20030073080A1
GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et
2.78e-92
888.50
89.22%
84.31%
86.09%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY: USA
ZIP: 63105
                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-917-563-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3834
                                                                                                    ٣
                                                                                                                                               15
                                                                                                                                                                                          34
                                                                                                                                                                                                                                                                                                                         94
                                                                                                                                                                                                                                                                                                                                                                     114
                                                                                                                                                                                                                                                                                                                                                                                                                                                           154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 194
  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
            Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                     ö
                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                               ò
                                                                                                                                                                   g
                                                                                                                                                                                          ò
                                                                                                                                                                                                             a
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                         ò
```

```
SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :::::|||{
3354 CGTAGGGGCCAGGAGATACTGCTTGGGCGAGGGCTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3414 TIGCIGCCCCATCACGCCTACGCCCAGCAGACGAGGCCTCCTAGGGIGTATAATC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9646
172
10
13
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 LysLysGlySerValValIleValGlyArgIleAsn--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-16 (1-197) x US-09-917-563-1 (1-9646)
                                                  CLASSIFICATION: ~......

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/238,076
FILING DATE: 26-7AN-1999
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REFERENCE/DOCKET NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4/
TELECOMMUNICATION INFORMATION:
TELEPHONE: 114-727-5188
TELEPHONE: 314-727-5188
TELEFAX: 314-727-5188
TELEFAX: 314-727-5188
TELEFAX: 314-727-5092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9646 base pairs
"WVR: NUCLEIC acid
             CURRENT APPLICATION DATA:
REPLICATION NUMBER: US/09/917,563
FILING DATE: 27-Jul-2001
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
AMTI SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
```

3473

3533

53

3713

3773

133

à g ò g

σ

3773

133

113

93

```
SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3834 TCGGGGGGTCCGCTGTTGTGCCCCGCGGACACGCCGTGGGCCTATTCAGGGCCGCGGGG 3893
                                                                                                                     74 AlaGlyThrArgThr1leAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                           94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 AlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 5, Application US/09995937
Publication No. US20030028010A1
GENERAL INFORMATION:
APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPAIITIS C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMPUTER: IBM PC compatible
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DSS/MS-DOS
SOFTWARE: Patentin Release $1.0, Version $1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 02.05/09/995,937
FILING DATE: 28-No. US20030028010A1-2001
CLASSIFICATION: <UNknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,756
FILING DATE: 04-May-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE CHARACTERISTICS:
LENGTH: 12980 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: ST. LOUIS
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NUMBER OF SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-995-937-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                             õ
                                                                                                                                           g
                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3414 ITGCTGGCGCCCATCACGGCGTACGCCCAGCAGACGAGGGCGTCCTAGGGTGTAAATC 3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                              CysthrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThrSerGlnThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                             Sequence 5. Application US/09238076
Patent No. US20020102540A1
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-16 (1-197) x US-09-238-076-5 (1-12980)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LysLysGlySerValVallleValGlyArgIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: HOWELL & HAFERKAMP, L.C. STREET: 7733 FORSYTH BLVD., SUITE 1400 CITY: ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6029-4831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORREY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
RECISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029
TELECOMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFANS: 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPARE: DATE: PC-DOS/WG-SOFTWARE: PAtrice PC-DOS/WG-DBD-COMPARE: PAtrice PC-DOS/WG-DBD-COMPARE: PAtrice Patri
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 314-727-6092 INFORMATION FOR SEQ ID NO: 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 12980 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.06e-92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               888.50
89.22%
84.31%
86.09%
                                                                                                                                                                             3954 ATGAGATCCCCG 3965
                                                                                                                     194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similari
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE: MCCOUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-238-076-5
                                                                                                                                                                                                                                                                      US-09-238-076-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ٣
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              34
   174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

ç

```
3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                 3414 TIGCIGGCGCCCAICACGGCGIACGCCCAGCAGACGAGAGGCCICCTAGGGIGIAIAAIC 3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        134 ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer 153
                                                                                                                                                                                                                                                                                                                                                     74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal 93
                                                                                                                                                                                          3 LysLysGlySerValValIleValGlyArgIleAsn-----------------------14
                                                                                                                                                                                                                                15 ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGlu 33
                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                       US-09-965-594-16 (1-197) x US-09-995-937-5 (1-12980)
                                                                                                                                                    Gaps:
                                                             Ś
                                                              ..
Q
                                                              SEQUENCE DESCRIPTION: SEQ ID
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHERICAL: NO
ANTI-SENSE: NO
                                                                                                  4.06e-92
888.50
89.22%
84.31%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11111111111
3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            194 MetArgSerPro 197
                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                           Alignment Scores:
                                                                       US-09-995-937-5
                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                          ò
                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

Search completed: August 31, 2003, 04:54:21 Job time : 190.482 secs

```
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Commun
```

Total number of hits satisfying chosen parameters: Minimum DB seq length: 0 Maximum DB seq length: 2000000000

22781392 seqs, 12152238056 residues

Searched:

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

em_estfun:* em_estom:* em_gss_hum:* em_gss_phg:* em_gss_vrl:* gb_gss1:* em_gss_inv:* em_gss_pln:* em_gss_vrt:* em_gss_fun:* em_gss_mam: ♣ em_gss_mus:* em_gss_pro:* em_gss_rod:* em_esthum:
em_esthum:
em_esthum:
em_esthu:
em_estpl:
em_htc:*
gb_estl:*
gb_htc:*
gb_estl:*
gb_estl:*
gb_estl:*
gb_estl:*
gb_estl:* em_estba:* EST:* Database :

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

n	
3	
-	
¥	
¢	
Σ	
3	
^	

	Description	BQ926101	BF631437 HVSMED001	2487	2274	6691	2842	545	1625	1 BJ02412	9	9	8		Ē	4	9	0	س ا	و ب	18 Danio	2 7	AT 555424	5 AGENC	1 60108	BI258851 602969666	zk29d0	5 21210	6 Singl	99	<u>``</u>	32 P	3 (7 :	ž i	ធី ខ	H 1	DON'S MGENCO	4625 CH240	TOTAL STORE STORE	33286 HOOTMO2	007 III-E-	6435 AL856435	BM825317 K-ESTO	
SUMMAKIES	110	BQ92610	143	9248	8227	0469	1284	9054	0162	302412	31617	95099	339	31	280	324	33	28	77	7 6	20.5	2 4	200	660960	E28991	BI258851	036834	E75761	CD	348	9	513	9 0	אור אור	4 (9 0	n •	4/600	2467	04660	8328	70800	6435	BM8253	
	h DB	3 13	9	9	9	-	2	٦ 1	2	 E	7	7	0	7	7	1	0	7		٦.	7.	-10	+ <i>o</i>	, כ	. 6	7 12	2	9	4	— ·		- i	٠. د د	⊣.	o (1 ·	ים מספ	, a	9 6		• •	, ,	. w		,
	Leng	14	7	11	αÓ	σ,	σ,	11	φ	φ	7	10	2	6	11	7	11	11	ω ;	12		7	* 0	13	9	75	2	S	æ	œ (σ,	9.	4.	4 4	n u	nc	, <u>.</u>	1	י ער	000	1	. ທ	S	, rv)
æ ,	Ouery	9.						•								•				٠	•	٠	•			8		•	٠	•	٠	٠	•	٠		٠	•		•	•					•
	Score	01.	٠	æ	86	6	97.5	97	96	96	96	96	95	95	95	97	94	σ	93.5	w, (ν (۰ ر	; ~	200	6	92	ä	ä	91.5	ä,	۲,	∹'	1,5	7.5	16	7.5	7.6	`		·	. თ	06	06	06	ı
	Result No.		~	٣	→	S	؛ ص	~	œ	σ	10	11	15	13	14	12	16	17	8 6	5 6	2 5	7 (3 6	2 5	25	56	27	28	59	30	31	7 6	, r	5 t	n (9 6	200	2 0	A		4.2		7 7	5	

ALIGNMENTS

RESULT 1 B0926101/c LOCUS DEFINITION ACCESSION VERSION	BO926101 AGENCOUT_8752655 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335718 5', mRNA sequence. BO926101. GI:22341132
AEI WORDS SOURCE ORGANISM REFERENCE	Mus musculus (house mouse) Mus musculus Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1403)

```
library
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                         EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Source
                                                                                                                             BF631437
LOCUS
                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1378 GGGTGTGTTCANCGGTCAGGACAGGTCGCC---GCACACTCGACGCGTCGGCCAGAGACT 1322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1111:::
1321 TGTCGGGGCGCGTTGGCGCATACCCCGGGTCGAGGTCAGCGCCGCTGTATACA 1262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1249 GGGGTA---TGGTTATCACGGCTGGGCAGGTACT------TCCCCTAAAGCG 1205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1204 GCGCCGTGCCGAGTATATATACCGGCGAGTGCCGAAGCGCACGGGGCGTGGAACGTTGACC 1145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .261 GAGGGGAAA-----CAG 1250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1087 GGCACGGTGATGGTTACTACGCGGCAGACCGGCTCCAACGGCGCTCTCTAACAGACGC 1028
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1. .1403
//Organism="Musmusculus"
//Organism="Musmusculus"
//Organism="Musmisculus"
//Organism="Musmisculus"
//Organism="Musmisculus"
//Iab_host="TuAGE:6335118"
//Iab_host="Dip="MiH_MGC_130"
//Orde="Organ: ologysts; Vector: pCMV-SPORT6.1.ccdb;
//Orde="Organ: ologysts; Vector: pcmv-Norte: pcmv-Norte
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               84 lyProvalThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProG 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                104 InGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrA 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              968
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65 GlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys--G 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11 GlyArgileAsnLeuSerGlyAspThrAlaTyrAlaGlnGLnThrArgGlyGluGluGly 30
                                                                                                                                Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
                                                                                                                                                                                                                       CDNA Library Preparation: ResGen, Invitrogen Corp

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bloscience Corporation

Clone distribution: MGC clone distribution information can be

found through the.I.M.A.G.E. Consortium/LLNL at:

http://lmage.llnl.gov

Plate: LLMAN3798 row: j column: 07

High quality sequence stop: 101.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31 CysGlnGluThr----SerGlnThrGlyArgAspLysAsnGln-----Val
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    45 GluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrCysIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         137 ----SerArgGlySer-----LeuLeuSerProArgProlleSerTyrLeuLysG
                                  National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1403
49
12
62
43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 124 rgHisAlaAspValIleProValArgArgArgGlyAsp----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                     Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-16 (1-197) x BQ926101 (1-1403)
      NIH-MGC http://mgc.nci.nih.gov/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15.5
101.50
36.75%
29.52%
9.84%
                                                                            Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        297
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

High quality sequence stop: 394. FEATURES Location/Qualifiers source 11199	/organism="Homo sapiens" /mol_type="mRNA"="Nome /db_xref="taxon:9606" /clone="IMAGE:6192708" /sex="male="type="sympathetic trunk" /tissue_type="sympathetic trunk" /dev_stage="adult, 16 yr" /lab_bost="blings" sympathetic trunk" /clone lib="Lubsk; sympathetic trunk"	/note="Vector: pCMV-SPORTE (Life Technologies); Site_1: Not1; Site_2: Sal1; cDNA made by oligo-dr priming. Directionally cloned using the following adaptors: 5'-TGACCCACGCGTCG-3' and 5'-GACTAGTTTAGATGGGGGGGGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.9 kb. This is a primary	11brary, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life Technologies. a 762 c 343 g 211 t 28 others ORIGIN	Alignment Scores: 24.8 Length: 1199 Pred No.: 98.50 Matches: 40 Score: 98.50 Matches: 40 Percent Similarity: 36.77% Conservative: 17 Best Local Similarity: 25.81% Mismatches: 54 Ouery Match: 9.54% Indels: 44 DB: 6	US-U9-965-594-16 (1-197) x BQB92487 (1-1199) OY 68 TrpThrValTyrHisGlyAhaGlyThrArgThrileAlaSerProLysGlyProValThr 87	108 SerLeuThProCysThr	Qy 119 LeuTyr-LeuValThr	Oy 143GlySerSerGl 155	Oy 175 rArgGlyValalaLysalavalAspPheIleProValGluSer 189 :::
BASE COUNT 147 a 216 c 311 g 105 t ORIGIN	Alignment Scores: 11.3 Length: 779 Pred. No.: 99.50 Matches: 47 Score: 9.50 Matches: 14 Best Local Similarity: 30.72% Mismatches: 63 Query Match: 10 Gaps: 6	-09-965-594-16 (1-197) x BF631437 (1-779) 49 GlnIleValSerThrAlaThrGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrp	Oy 69 ThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr 87	Qy 102 AlaProdInGlySerArgSer		OY 165 laValGlyllePheArgAlaAlaValCysThrArg 176	HO892487 LOCUS LOCUS LOCUS BQ892487 LOCUS LOCUS BQ892487 LUPSK1_Sympathetic_trunk Homo sapiens cDNA clone IMAGE:6192708 5', mRNA sequence. ACCESSION BQ892487.1 GI:22284501 KEYWORDS EST. FOUNCE Homo sapiens (human)	ORGANISM Homo saplens Eukkaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Bukaryota; Metazoa; Chordata; Catarrhini; Hominidae; Homo. REFERENCE I (bases 1 to 1199) AUTHORS NIH-MGC http://mgc.nci.nih.gov/. TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished COMMENT Contact: Robert Strausherg, Ph.D.	Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. James R. Lupski cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LiNL at: http://image.llnl.gov Plate: LLAM13595 row: c column: 13 High quality sequence start: 57

us-09-965-594-16.rst

```
BF304699.1 GI:11251586
                                                                                                                                                                                  EST.
Homo sapiens (human)
                   176 gGlyValAlaLys 180
                                        21.4
98.00
44.058
38.108
9.50%
                                                                                                                                                                                                                                                                                                           Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                             RESULT 5
BF304699/c
LOCUS
                                                                                                                        DEFINITION
                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
                                                                                                                                                                                                                                                             REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                       ACCESSION
                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                             Email: cgapbs_r@mail.hib.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The Life Technologies, Inc.
cDNA Library Arrayed by: The LiA.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://mage.lnl.gov
Plate: LLAM9308 row: g column: 07
High quality sequence stop: 696.
Coation/Qualifiers
1. 846
/ organism="Mus musculus"
/ drain="C57BL/6J"
/ drain="TRAGE: 4035102"
/ tissue=type="mumor; gross tissue"
/ dew_stage="mumor; gross tissue"
/ dew_stage="mumor; gross tissue"
/ dew_stage="mumor; gross tissue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               122 IThrargHisalaAspvallleProValArgArgArgGlyAspSerArgGlySerLeuLe 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       uSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeu---LeuCysPr 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I bases 1 to 846; NIS. NIH-WGC http://mgc.nci.nih.gov/.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /note="Organ: mammary; vector: pCMV-SPORT6; Site_1: Sall; Site_2: Not1; Cloned unidirectionally. Primer: Oligo dT. Library constructed unidirectionally primer: providing samples: Lothar Hennighausen/Robin Humphreys, providing samples: Lothar Hennighausen/Robin Humphreys.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----AlaProG1 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   nGlySerArg-----SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVa 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    161 oAlaGlyHis---AlaValGlyIlePheArg------AlaAlaValCysThrAr 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyAlaGlyThrArg-ThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAs 92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92 nValAspLysAspLeuValGlyTrpGln--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          846
46
14
35
31
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /clone_lib="NCI_CGAP_Mam5"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                210
                                                                                                                                                    Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-16 (1-197) x BF182274 (1-846)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241 g
                                musculus (house mouse)
musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             218 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17.6
98.00
48.00%
36.80%
9.50%
                                                                                                                                     Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     555
                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                              ORGANISM
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             \ddot{o}
```

```
1. .984

(organism="Homo sapiens"

/mol_type="mana"

/mol_type="mana"

/mol_type="mana"

/db_xref="taxon:9606"

/clone="langle="tabdomyosarcoma"

/lab_host="bhi08 (phage-resistant)"

/clone=lub="NHH_MGCl7"

/note="Organ: muscle; vector: poTB7; Site_1: EcoRI;

/note="Organ: muscle; vector: poTB7;

/note="Organ: muscle; vector: po
984 bp mRNA linear EST 21-NOV-2000
601888252F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4122276 5',
mRNA sequence,
BF304699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100 TrpGlnAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota: Metazoa: Chordata: Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates: Catarrhini; Hominidae; Homo. 1 (bases 1 to 984) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Robert Strausberg, Ph.D.

Contact: Robert Strausberg, Ph.D.

Email: cgapbs - remail.nih.gov

Lissue Procurement: ArCc

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Preparation: Ling Hong/Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incytte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: LLCM1005 row; g column: 13

High quality sequence stop: 646.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          119 LeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  984
32
52
52
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-16 (1-197) x BF304699 (1-984)
```

MEDLINE PUBMED REFERENCE AUTHORS

JOURNAL

ø

```
/protein_id="BAG37940.1"
/db_xref="G1:26348601"
/db_xref="G1:26348601"
/translation="bSCIPEAASPERSLITPRGDGFFLKEKLSAARAVGPGPSVAFGVS
TKYRGAQAGQGRVGACRRSEGICLSRKPRRQHVPVGPHYYGLSGGRRIPPPAGE
AQAAGBAPQQVPHPGRRHGTVVPPQGAAGLLPALAARQVPVGPVRGFREGPRRAPRHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BJ001625
BJ001625 MF01SSA CDNA Oryzias latipes CDNA clone MF01SSA025C02 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Oryzias latipes (Japanese medaka)
Oryzias latipes
Cyzias Altipes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120 rLeuValThrArgHisAlaAspValIleProValArg---ArgArgGlyAspSerArgGl 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                100 pGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            411 -credenderedecantecerations and the contraded contra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              172 avalcysthrargGlyvalalauysalaval-----aspPhelleProvalGluSerLe 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81 Ser-ProLysGlyProvalThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTr 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 ySerLeuLeuSerProArgProlle------SerTyrLeuLysGlySe 153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    153 rSerGlyGlyProLeuLeuCysProAla -- - GlyHisAlaValGlyIlePheArgAlaAl 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            587 AGTGTGTGAGTGCGGAGACTTGAGGGTGGTTGGCTCTGAGCTGACCTGAAGCTCCTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [Mus
                                               /note="unnamed protein product; putative weakly similar to zinc finger protein (fragment) musculus] (PIR[148722, evidence: FASTY, 50.7%ID, 57.6%length, match-601)" /codon_start="
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1141
111
113
33
35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ل
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TrpThrValTyrHisGlyAlaGly----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-16 (1-197) x AK080545 (1-1141)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             353
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note="putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note="putative"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BJ001625.1 GI:17364516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1118. .1123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190 uGluThrThrMet 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               647 GGCAGGTACACTT 659
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32.5
97.00
39.318
31.728
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             316 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 mRNA sequence.
BJ001625
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          polyA_signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              polyA_site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
BJ001625/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VERSION
KEYWORDS
SOURCE
          CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                   Kawal, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishli, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Arawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Caito, T., Caito, T., Caito, T., Caito, T., Caito, T., Saito, T., Radota, K., Matsuda, Y., Gissi, C., King, B., Kochiwa, H., Kuchl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Kochiwa, H., Kuchl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Tomita, M., Magner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Baldarelli, R., Bromstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Toyo-oka, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wanshizaki, Y., and Hayashizaki, Y., Rasagawa, Y., Kawaji, H., Kohtsuki, S., Matsuki, Y., Kawaji, H., Kohtsuki, Y., Kawaji, H., Kohtsuki, Y., Ramani, H., Kawaji, H., Kohtsuki, Y., Rawaji, H., Kohtsuki, Y., Kawaji, H., Kawaji, H., Kawaji, H., Kawamana, K., Kawaji, H., Kawaji, H., Kawaji, H., Kawamana, K., K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            L Nature 420, 563-573 (2002)

E (bases 1 to 1141)

S Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, R., Hiraoka, T., Hirozane, T., Hayashida, K., Hayatsu, N., Hiramoto, R., Hiraoka, T., Hirozane, T., Katch, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Ohasto, N., Okazaki, Y., Saito, R., Nomura, Y., Nishi, K., Numuraka, R., Numazaki, R., Oho, M., Ohasto, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, C., Sakai, T., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Soqabe, Y., Taqami, M., Tagawa, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y., Toya, T., Yasunishi, A., Muramatsu, M., And Mayashizaki, Y., Toya, T., Yasunishi, A., Muramatsu, M., Murama
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Submitted (16-ARR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Physical and Chemical Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute: 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-reségsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. DIVISION of Experimental Animal Research in Riken contributed to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The FANTOM Consortium and the RIKEN Genome Exploration Research
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               prepare mouse tissues.
Please visit our web site for further details
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /moi_type="mRNA"
/strain="C57BL/6"
/db_xref="FaNTOW DB: N730082L10"
/db_xref="taxon:10090"
     Res. 10 (11), 1757-1771 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              URL:http://genome.gsc.riken.go.jp/
URL:http://fantcom.gsc.riken.go.jp/
Location/qualiflers
1. .1141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /clone-"A730082L10"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pax:81-45-503-9216)
Genome Re
20530913
11076861
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11217851
```

JOURNAL REFERENCE AUTHORS

TITLE

TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL

COMMENT

FEATURES

```
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neoplerygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha; Beloniformes; Acinichthyidae; Oryzlinae; Oryzlas.

( Dases 1 to 643)
Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H., Medaka Est Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       96 AspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGly 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Arg-----ProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    242 AAAAATGACGTAGAACCAAAAGACACAGATCCAACCACACATGTTCTGGTTCTACGGGCT 301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             344 GCTCTGCGGCGGCGAGGAGCTCCTGGGCCAGTTGTGACTCGT------385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----GGAGGACGAAGAGCGTCACCCCGGAGCTGTAGGCTGCAGGGATGCGGATGTGGC 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SerSerAspLeuTyrLeuValThrArg------124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            440 TCTGCT-----TTGGTTCCTCGTCCTCGATCATCTTCCTCACCTGACCTTCCA 490
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    125 HisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLeuSerPro 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         491 CATCCAGGIGICGCCAGCGCIGGTCIGACGGGIGAIGGGAGAGGCCGGACAGGCAGCAGT 550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       551 CGGGGTGAATCTCTGCAGGACGTCTTCACGGCGGATCAGGAGGACCGACTCGCTGCAGA 610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 ThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81 SerProLys------GlyProValThrGlnMetTyrThrAsnValAspLys 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /clone="MF01SSA143D12"
/sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/clone_lib="MF01SSA cDNA"
148 c 148 g 176 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               643
7
7
58
36
6
                                                                                                                                                                                                                                                                                                    Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mibhima, Shizuoka 411-8540, Japan
Tat: 81-559-81-6856
Fax: 81-559-81-6855
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:

    643
    /organism="Oryzias latipes"
/mol_type="mRNA"
    /strain="Hd-rR"

                                                                               Oryzias latipes (Japanese medaka)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-16 (1-197) x BJ024121 (1-643)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /db_xref-"taxon:8090"
                                                                                                                                                                                                                                                                                                                                                                                                       Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
                       BJ024121
BJ024121.1 GI:17377389
                                                                                                                                                                                                                                                               Unpublished
Contact: Tadasu Shin-i
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19.4
96.00
34.72%
29.86%
9.30%
                                                                                                      Oryzias latipes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                              EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No.:
                    ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BJ024121 BJ024121 MF01SSA CDNA OFYZIAS latipes CDNA clone MF01SSA143D12 3',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -------TGTTGGAGAACCTATCACAGTTCCTGCTTTAGAGCAACGGCA 410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96 AspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGly 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  111:::
313 TCTGCT------TTGGTTCCTCGTCCTCGGATCATCTTCCTCACCTGACCTTCCA 263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeuLeuSerPro 144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SerSerAspLeuTyrLeuValThrArg----- 124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          262 CATCCAGGTGTCGCCAGCGCTGGTCTGACGGGTGATGGGAGGGCCGGACAGGCAGCAGT 203
                                   1 (bases 1 to 615)
Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
Medaka EST Project in Takeda's lab
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 ThrCyslleAsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrlleAla 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81 SerProLys------GlyProvalThrGlnMetTyrThrAsnValAspLys 95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----GGAGGACGAAGAGCGTCACCCCGGAGCTGTAGGCTGCAGGGATGCGGATGTGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
                                                                                                                                                                                                                                                                                                                                                                                                   /sex="mixture of female and male"
/tissue_type="whole embryo"
/dev_stage="segmentation stage 20 - 25"
/clone_lib="MFOLSSA cDNA"
166 c 165 g 144 t
                                                                                              Unpublished
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yate, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Fax: 81-559-81-6855
Email: tshiniqgenes.nig.ac.jp.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  615
43
7
58
36
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                  /organism="Oryzias latipes"
/mol_type="mRNA"
/strain="Hd-rR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-16 (1-197) x BJ001625 (1-615)
                                                                                                                                                                                                                                                                                                                                                             /db_xref="taxon:8090"
/clone="MF01SSA025C02"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18.3
96.00
34.72%
29.86%
9.30%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  162 AlaGlyHisAla 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GCCTCTGCTGCA 131
                                                                                                                                                                                                                                                                             1, .615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             367
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 9
BJ024121
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment S
Pred. No.:
                                       REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BASE COUNT
                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGIN
```

g

ò 셤 õ g

à

õ 셤 ô a ò g ò 셤 õ

0y 125 HisalaAspvalIleProValArgArgArgArgGLyAspSerArgGlySerLeuLeuSerPro 144 1 1 1 1 1 1 1 1 1 1	RESULT 11 CB950999 LOCUS LOCUS DEFINITION AGENCOURT_13445496 NIH_MGC_177 Mus musculus cDNA clone IMAGE:30316162 5', mRNA sequence. ACCESSION CB950999.1 GI:30205777 KEYWORDS FOR Mus musculus (house mouse) SOURCE NGGANISM Mus musculus CRGANISM Mus musculus DEFATOR (house mouse) ORGANISM Mus musculus Amamalia; Butheria; Rodentia; Sciurognathi; Muridae; Mus. Mus.	REFERENCE 1 (bases 1 to 1031) AUTHORS NIH-MGC http://mgc.nci.nih.gov/. AUTHUE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Contact: Robert Strausberg, Ph.D. Email: cgapbs-rémail.nih.gov Tissue Procurement: Dr. Michael Brownstein cDNA Library Preparation: Michael Brownstein Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://mage.llnl.gov Plate: NDCM107 row: b column: 11	High quality sequence stop: 333. FEATURES Location/Qualifiers 11031 /organism="Mish musculus" /mol_type="MRNA" masculus" /mol_type="MRNA" masculus" /db_xref="taxon:10090" /clone="iMASE:30316162" /lab_host="DH10B (11-phage-resistant)" /clone=lib="MH40B (11-phage-resistant)" /clon	Adaptors were used in cloning as follows: 5AAGCAGGGGAGAGGCGAGAGGGCGGGGGGGGGGGGGGG	Alignment Scores:
16 61 0N		oan	### Addition stage 20 - 25" ### Addition stage 20 - 25" ### Clone_lib="MF0ISSA CDNA" ### Construction stage 20 - 25" ### Clone_lib="MF0ISSA CDNA" ### Alignment Scores: ### Align	-09-965-594-16 (1-197) x BJ016176 (1-754) 41 LysAsnGlnValGluGlyGluValGlnIleValSerThralaThrGlnThrPheLeuAla	0y 81 SerProLys

US-09-965-594-16 (1-197) x CA728398 (1-580)	BF203316 601865914, mRNA sequ BF203316, BF2033	/organism="Homo sapiens" //organism="Homo sapiens" //mol_type="manna" //db_xxef="taxon:9606" //clone="InAGE:4098578" //tissue_type="InAddomyosarcoma" //tissue_type="InAddomyosarcoma" //tissue_type="InAddomyosarcoma" //tissue_type="InAddomyosarcoma" //tissue_type="InAddomyosarcoma" //tissue_type="InAddomyosarcoma" //clone_lip="NIH MGC_l7" //clon
0y 64 AsnGlyValCysTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys 83 1	CA728398 N wdllc.pk0 end, mRNA CA728398 CA728198 EST. Triticum ENAryota Spermarop ; Triticum Cupublish Cupublish Contact: Crop Gene E. I. Dup I. Innovat: Crop Gene E. I. Dup I. Innovat: Tel: 302- Fax: 302-	FEATURES Location/Qualifiers source location/Qualifiers /organism="Triticum aestivum" /nol_type="mRNA" /db_xref="taxon:4565" /clone="wdit.ppe="inflorescence" /lissue_type="inflorescence" /lish_bost="bh108" /clone_lib="wditc" /note="wector: pBluescript SK+; Site_1: EcoRI; Site_2:

Percent Similarity: 34.62% Conservative: 18 Best Local Similarity: 24.73% Mismatches: 62 Query Match: 9.21% Indels: 57 DB: 12 Gaps: 10 US-09-965-594-16 (1-197) x BM915803 (1-1146) Oy 22 AlaginGlnThrArgGlyGluGluGlyCysGlnGluThrSer-GlnThrGlyArgAspLy 41	1002 CGCCGCTGGTGGGATGTTGTGG 77	y 107 gSetLeuthrProCysThrCysGlySetSerAspLeuTyrLeuValThtArgHisAlaAs 127 1858 ACTTGCTTGTCGTGTTTTCTGTGGG	16 74 74 78 78 79 79 79 79 79 79 79 79 79 79 79 79 79	Lefebvre, P., McDermott, J.P., Shrager, J., Silflow, C. and Stern, D. Analyses of the Chlamydomonas reinhardtil Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants: project phase 3 JOURNAL Contact: Charles Hauser COMMENT Contact: Charles Hauser COMMENT DAB Box 91000 Duke University Durham, NC 27708-1000 Tel: 919 613 8177 Fax: 919 613 8177 Email: chauser@duke.edu.
133 553 153 493 173	Db Qy BST 12-MAR-2002 Qy Db	Db D	n be Site_2: Iv cloned	daptor: in the alifornia, gene) and his is a
	rArgGlyvalAlaLys 180 	5.7, mRNA sequence. BM915803.1 GI:19366182 BM915803.1 GI:19366182 EST. Homo sapiens (human) Homo sapiens Homo	Contact: Robert Strausberg, Ph.D. The sense of the sense	into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGGG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of Californ) Berkeley) using 2AP-cDNA synthesis kit (Stratagene) a Superscript II RT (Life Technologies). Note: this is NIH MGC Library. 492 c 344 g 141 t 51.3 Length: 1146 95.00 Matches: 45
Best Local Similarity: 40.58% Query Match: 9.21% DB: 10 US-09-965-594-16 (1-197) x BF203316 Qy 114 CysGlySerSer-AspleuTy:	173 alcysthu 1 1 1 1 1 1 1 1 1 1	ACCESSION BM915803 VERSION BM915803 I GI:19366182 KEYWORDS EST. SOURCE HOMO SAPIENS (human) ORGANISM HOMO SAPIENS (human) Manualia; Eutheria: Pri REFERENCE I (bases I to 1146) AUTHORS NIH-MGC http://mgc.nci.TITLE National Institutes of JORNAL Unpublished	COMMENT CONTACT: Robert Strausberg, Ph Email: cgapbys refmail.nih.gov Tissue Procurement: DCTD/DTP CDNA Library Preparation: Rub: CDNA Library Preparation: Rub: CDNA Comment of Straibution: MGC clone found through the I.M.A.G.E. CA http://image.lin.gov Plate: LLCM2007 row: i column High quality sequence start: 6 High quality sequence start	into EcoR: GGCAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG

```
1...'Organism="Chlamydomonas reinhardtii"

//organism="Chlamydomonas reinhardtii"
//mol_type="mRNA"
//strain="Cc-1690 wild type mt+ 21gr"
//db_xref="taxon:3055"
//clone_lib="C. reinhardtii CC-1690, Stress condition I,
normalized, Lambda Zap II"
//note="Vector: pBluescript II SK-; Site_l: ECORI; Site_2:
//note="Vector: pBluescript II SK-; Site_l: ECORI; Site_2:
//note="Vector: pBluescript II SK-; Site_l: ECORI; Site_2:
//note="Vector: pBluescript II SK-; Site_1: ECORI; Site_2:
//note="Vector: pBluescript II SK-; Site_1: Ahr), TAP-F (30 min,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr, 24hr), NO3 to NH4 (30min, Ihr,
Ihr, 4hr), TAP-P (4hr, 12hr,
Ihr, 4hr), TAP-P (4hr, 12hr,
Ihr, 4hr), TAP-P (4hr, 12hr,
Ihr, 4hr), TAP-P (30min, Ihr,
Ihr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87 ThrGlnMetfyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           231 ACAAAGTTACATACACGG-----AAGGACCAGCGCGTTGGCCACCCCCTTGGAGCCG 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              107 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           285 AGAAGCCCGACCGCTGGTCCTCTGGGGTCATCCGCATGCTATGCAATCTCCCGGTATCAG 344
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          127 Aspvalile ----- 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----GlySerLeuLeuSerProArgProIleSerTyrLeu 150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             405 GAAGGGGACGGGAATTCCGTTTGCGGAAAAGTGAGCGCCCCAAGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 345 GAGATCATITGCATGTGGCTTTAGTCACCCCAAGAGAGCCTGGGAGTGGGCATTTATAA 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                71 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys-----GlyProVal 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             701
32
14
45
22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                151 LysGlySerSerGlyGlyProLeuLeuCysProAlaGly 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  165 CTCCAAGGCAGCAATGGGAGCCTTTCGCGGTGTGCGGGT 503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-16 (1-197) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Search completed: August 31, 2003, 04:27:34
Job time : 1916.31 secs
       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             33.9
94.00
40.71%
28.32%
9.11%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
ORIGIN
                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
```

```
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Z K Z E
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                            (without alignments)
700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1: /SIDSI/goddata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/goddata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/goddata/geneseq-gembl/AA1981.DAT:*
4: /SIDSI/goddata/geneseq-gemenbl/AA1981.DAT:*
5: SIDSI/goddata/geneseq-gemenbl/AA1982.DAT:*
6: /SIDSI/goddata/geneseq-gemenbl/AA1982.DAT:*
7: /SIDSI/goddata/geneseq-gembl/AA1980.DAT:*
8: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1980.DAT:*
8: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1980.DAT:*
9: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1980.DAT:*
10: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1990.DAT:*
11: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1991.DAT:*
12: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1991.DAT:*
13: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1991.DAT:*
14: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1991.DAT:*
15: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1991.DAT:*
16: /SIDSI/goddata/geneseq-gemeseqp-embl/AA1991.DAT:*
17: /SIDSI/goddata/geneseq-gembl/AA1991.DAT:*
18: /SIDSI/goddata/geneseq-gembl/AA1999.DAT:*
18: /SIDSI/goddata/geneseq-gembl/AA1999.DAT:*
18: /SIDSI/goddata/geneseq-gembl/AA1999.DAT:*
18: /SIDSI/goddata/geneseq-gemebl/AA1999.DAT:*
18: /SIDSI/goddata/geneseq-gembl/AA1999.DAT:*
20: /SIDSI/goddata/geneseq-gembl/AA1999.DAT:*
21: /SIDSI/goddata/geneseq-gembl/AA1999.DAT:*
22: /SIDSI/goddata/geneseq-gembl/AA2001.DAT:*
23: /SIDSI/goddata/geneseq-gembl/AA2001.DAT:*
24: /SIDSI/goddata/geneseq-gembl/AA2001.DAT:*
24: /SIDSI/goddata/geneseq-gembl/AA2001.DAT:*
                                                                                                                                                                                                                     ......VAKAVDFIPVESLETIMRSP 197
                                                                                                          August 30, 2003, 17:42:58 ; Search time 44.6227 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                   1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                               1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                       Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                       protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15223
AAB15224
AAB15222
AAB15222
AAB15221
AAB15220
AAB15220
AAB15226
                                                                                                                                                                                                                                                        BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                              US-09-965-594-18
1017
1 MKKKGSVVIVGRINLSCDTA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ü
                                                                                                                                                                                                                                                                                                                                                                                    length: 0
length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              197
197
197
195
195
665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0
98.8
98.5
97.8
93.5
93.0
89.7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Title:
Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1017
1005
1005
995
985
951
946
                                                                                                                                                                                                                                                        Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq
Maximum DB seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                                                       OM protein
                                                                                                                                                                                                                                                                                                               Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database
                                                                                                            Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     E 450 6 7 8 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
```

-		75.1	9	4.	665	20	AAY 24947		HCV	NS4A-NS	3 compl
7,7	- ~	v ri	o o	n m	665 665	202	AAY24941 AAY24942		HC 4	NS4A-NS NS4A-NS	3 compl
-	ω.	5.1	9	0	216	50	AAY17880		HCV	NS4A-NS	3 compl
77	4 N	v. r.	99	00	665 665	70 70	AAY24945 AAY24946		HC A	NS4A-NS	3 compl
	9.0	ω. r.	S	σ,	665	50	AAY24940		HC.	NS4A-NS	3 comp1
17	8 871.	. r.	85.	٠. <i>د</i>	216	50	AAY 24 94 8 AAY 17884		HCV	HCV NS4A-NS3 HCV NS4A-NS3	3 compl
- 10	o С	. י	S	vo vo	216	20	AAY17879		HCV	NS4A-NS	3 compl
		. 2.	S	9	665	20	AAY24944		HCV	NS4A-NS	s comp1
	N 10	370	2	e G	671 215	202	AAY 24949		HCV	NS4A-NS	3 compl
	4	.5.	2	m	216	20	AAY17882		HCV	NS4A-NS	3 compl
	n v	v. r.	S O	m r	216	200	AAY17883 AAY17886		HCV	NS4A-NS	3 compl
	, _	. 2.	2	(1	216	22	AAY17877		HCV	NS4A-NS	3 compl
.,	∞ c	364	5	0 0	215	50	AAY17887		HCV	NS4A-NS	3 compl
• •	, C		* 4	ם ע	216	2 5	AAY1/861 AAY17885		2 2	NS4A-NS	3 comp1
) (P)		929	4	'n	213	20	AAY17888		HG A	NS4A-NS	comp1
mir	01	959	ͺ	ın -	631	200	AAW93482		HCV	NS3 pro	tein.
, (স ৰ	ų v	4 4	* ₹	191	7.5	AAY44/28	•	Hepa	titis C	virus
1 (*)	. 10		* 4	, 4	3011	4.4	ABP71460		Amino	o acid	Seguenc
	9	5.5	毋 .	~ ,	3012	23	AAU99289		Hepa	itis	c virus
., (,	~ a		* *	40	3011	4 4	AAR40120		HCV	genomic	amino
, (*)	oon		. 4	0	1648	16	AAR79221		DHCV176	176-encoded	7 P
•	0	15	4	0	1766	10	AAP92041		Sedu	•	
•	- (i.	*	0 0	1786	2:	AAP90158		Protein	Ø)	8,
7 7	7 M	u ru	4 4	5 C	2301	25	AAP90164 AAP92047		Secuenc	a)	ਹ 4
•	4	.5	4	0	2436	10	AAP92050		Segu	, e	e.
4	ς.	5.	4	0	2436	10	AAP90288		Peptid	e en	73
							ALIGNMENT	TS			
RESULT AAB152	1 23 88152	23 643	Ti		40		44 50				
×	•	,					;				
S S	AAB152	23;									
מלט	19-DEC	-2000	(f	irst	entry	(¥					
XX DE	Hebatil	tisc	V	rus	NS4A-NS	4	usion protease	5# 986			
×))	•			1		•			
X X X	Hepati liver	:18; [a1]u	NS3	prot live	otease; ver can	vir Ger;	al replic mutant;	; chronic n.	liver d	disease;	
S S S	Hepatiti Syntheti	tis c	Vir	. sn.							
Y N	WO2000407	07	-A1.								
×	13-лпг.	-2000									
X			•	CETO	Č	•					
. ×	2		7) } }	0000-0000	.245.					
PR S	08-JAN	-1999		sn66	-0115	271.					
V & S	(BRIM) BRI	STOL	STOL-MYERS		SQUIBB	.00				
J.	Wittek	ekind M		einh	Weinheimer	ŝ	Zhang Y,	Goldfarb V;			
X K	WPI; 20	\$ -000	629	76/40	_•						
X X	-PSDB	¥.	733								

N

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 ISINGVLWIVYHGAGTRIJASPRGPVIQMITNVDKDLVGWQAPQGGRSLIPCICGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 I.VTRHADVIPVRRRCDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                          substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                          1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15224 standard; protein; 197
                                                                 Claim 23; Fig 15; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity
Matches 197; Conserv
                                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15224;
                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                         엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
```

Ş

(first entry)

Weinheimer S, Zhang Y, Goldfarb V;

990S-0115271

ö The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-7 61 TSINGVLWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120 121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180 9 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA Hepatitis, NS3 protease, viral replication, chronic liver disease, liver failure, liver cancer; mutant; mutein. ö Score 1005; DB 21; Length 197; Pred. No. 2.7e-95; 2; Mismatches 1; Indels 0 Goldfarb V; Hepatitis C virus NS4A-NS3 fusion protease #4 Wittekind M, Weinheimer S, Zhang Y, AAB15222 standard; protein; 197 AA 8 66pp; English 181 AVDFIPVESLETTMRSP 197 98.8%; 98.5%; (BRIM) BRISTOL-MYERS SQUIBB 06-JAN-2000; 2000WO-US00345 990S-0115271 (first entry) Best Local Similarity 98.5 Matches 194; Conservative WPI; 2000-465976/40. Hepatitis C virus. Synthetic. 197 AA; Claim 23; Fig 16; N-PSDB; AAA73331 WO200040707-A1. 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Seguence AAB15222 Query Match Best Local S AAB15222 g QQ a g ò ò ò

ő

100.0%; Score 1017; DB 21; Length 197; 100.0%; Pred. No. 1.6e-96; 1ve 0; Mismatches 0; Indels 0;

Conservative

~

```
Matches 194; Conservative
                                                                                                                                                                                                                WPI; 2000-465976/40
                                                                                                                                                                    Hepatitis C virus.
Synthetic.
                                                                                                                                                                             WO200040707-A1
                                                                                                                                                                                                           Wittekind M,
                                                                                                                                                                                               38-JAN-1999;
                                                                                                                                                                                   13-JUL-2000
                                                                                               61
                                                                                                          121
                                                                                                                181
                                                                                                                      181
                                                                                                                                           AAB15225;
                                                                                         61
                                                             Sequence
                                                                  Query Match
Best Local
                                                       variant
                                                                                                                               RESULT 4
                                                                                                                                        දු
                                                                                         ò
                                                                                               ద
                                                                                                     ŝ
                                                                                                          a
                                                                                                                      ద
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to claave lits replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helixo-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                       substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophillic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                      Modified hepatitls C virus (HCV) NS3 protease comprising at least 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 995; DB 21; Length 197;
Pred. No. 2.9e-94;
2; Mismatches 2; Indels C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ş
                                                                                                                                       Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15221 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVDEIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97.84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       tches 193; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
N-PSDB; AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                    The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                    substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 1002; DB 21;
Pred. No. 5.6e-95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease #7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15225 standard; protein; 197
                                                                                                                    Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVDFIPVESLETTWRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                         98.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                     197 AA;
```

ö

Gaps

ö

9 9

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helixo-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TSINGVLWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                      Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                        ë.
                                                                                                                                                                                                                                                                                                                                                                                       Score 985; DB 21;
Pred. No. 3.2e-93;
0; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                           Claim 23; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15220 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity 97.0%;
Matches 191; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
2000-465976/40
                                                                                                                                                                                                                                                                                                                                                          197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
Synthetic.
WPI: 2000-465976,
N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15220;
                                                                                                                                                                                                                                                                                                                               /ariant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      合
```

```
ä
                                                                                                                                                              The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28
                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TSINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDRDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                         ..
                                                                                                                                                                                                                                                                                                                                                                                                    Score 951; DB 21; Length 195;
Pred. No. 1e-89;
1; Mismatches 7; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                    Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB15226 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                      Ouery Match
Best Local Similarity 94.9%;
Matches 187; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                          195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                            N-PSDB; AAA73329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15226;
                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                            variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 7
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
```

ö

Gaps

;; 0

n

```
28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                              Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                       ö
                                                                                                                       The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                   TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                            LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                          1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis, NS3 protease, viral replication; chronic liver disease;
liver failure, liver cancer.
                                                                                                                                                                                                                                                                                 Length 197;
                                                                                                                                                                                                                                                                                                      Indels
 Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                 Score 946; DB 21;
Pred. No. 3.3e-89;
0; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15212 standard; protein; 195 AA
                                                                                                     Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                93.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   06-JAN-2000; 2000WO-US00345
 ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                           Best Local Similarity 94.4
Matches 186; Conservative
 Weinheimer
                   WPI; 2000-465976/40.
N-PSDB; AAA73335.
                                                                                                                                                                                                                                                             197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WO200040707-A1.
 Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15212;
                                                                                                                                                                                                                                                                                 Query Match
à
                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                      đ
                                                                                                                                                                                                                                                                                                                                                                                                                             đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                           ð
```

```
LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAR 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                 The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the Polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MKKKGSVVIVGRINLSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                .;
;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 912; DB 21; Length 195;
Pred. No. 1e-85;
Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1; Mismatches
Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ş
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                  Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY24943 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AVDFIPVESLETIMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 89.7%;
92.4%;
Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-0S24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0094331,
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETTMRSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 182; Conservative
                                                        2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                     N-PSDB; AAA73328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY24943;
```

ή

z

```
VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                                                                                                       The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 estine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 86.7%; Score 881.5; DB 20; Length 665; Best Local Similarity 85.2%; Pred. No. 8e-82; Matches 167; Conservative 16; Mismatches 10; Indels 3;
                                                                                                                        New hepatitis C virus covalent complexes
                                                    Yao
                                                                                                                                                           Claim 6; Page 90-92; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY24947 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex SEQ ID NO:18.
                                                    PC,
                                                      Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                      Taremi SS,
                  (SCHE ) SCHERING CORP.
                                                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                          665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-SEP-1999
                                                      Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY24947;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            202
                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24947
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

```
62 SINGVLWITVYHGAGIRTIASPRGPVTOMYINVDRDLVGWQAPQGSRSLIPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82 CVNGVCWTVYHGAGSKTLAGPKGPITOMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSAGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                                                                                                  NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                       The present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .;
m
                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                               / Match
Local Similarity 84.7%; Pred. No. 1.6e-81;
les 166; Conservative 17; Mismatches 10; Indels
                                   ż
                                                                                                     New hepatitis C virus covalent complexes
                                                                                                                                        Claim 6; Page 100-102; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ź
                                   Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV NS4A-NS3 complex SEQ ID NO:12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24941 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US24528.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-SEP-1999 (first entry)
                                   Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SCHE ) SCHERING CORP
(SCHE ) SCHERING CORP
                                                                  WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                   665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-NOV-199B;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-1998;
28-NOV-1997;
                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24941;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182
                                                                                                                                                                                                                                                                                                                                                                                                   Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3; Gaps
```

```
122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                                                                                                                                                                                                            NSAA PERIOR COMPISSION CENTRAL HOLDS AND COMPLEX COMPL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKITSLTGRDKNQVBGEVQVVSTATQSFLAT 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                         present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV; hepatitis C virus; single chain recombinant complex; linker;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 86.3%; Score 877.5; DB 20; Length 665; Best Local Similarity 85.2%; Pred. No. 2.1e-81; Matches 167; Conservative 15; Mismatches 11; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15; Mismatches
                                                                                                                                                 New hepatitis C virus covalent complexes
                                                                                                                                                                                                             Claim 6; Page 85-87; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY24942 standard; Protein; 665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complex SEQ ID NO:13.
                            5
                            Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07-SEP-1999 (first entry)
                            Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SCHE ) SCHERING CORP.
                                                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24942;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY24942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
```

```
62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                     NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 peptide. A linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 peride is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                                                         22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVBGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                         5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                       present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                               the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                        .;
                                                                                                                                                                                                                                                                                                                                                            DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                         11; Indels
                                                                                                                                                                                                                                                                                                                                                         86.3%; Score 877.5; DB 2
85.2%; Pred. No. 2.1e-81;
ative 15; Mismatches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Yao N;
ż
                                                           New hepatitis C virus covalent complexes
                                                                                          Claim 6; Page 88-90; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ş
P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY17880 standard; Protein; 216
Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    970S-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0094331
                                                                                                                                                                                                                                                                                                                                                                         Local Similarity 85.49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Malcolm BA, Taremi SS,
Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP
                             WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                             665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                              Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY17880;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182
                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY17880
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

```
62 SINGVLWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 estrine protease domain. Where the hydrophobic domain of native HCV NS3 serine protease domain. Where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Appase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                              5 GSVVIVGRINLSGD -- - TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                          22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                  3;
                                                                                                                                                                                                                                                                                                                                                  Length 216;
                                                                                                                                                                                                                                                                                                                                                                                  10; Indels
                                                                                                                                                                                                                                                                                                                                                  86.0%; Score 874.5; DB 20;
85.1%; Pred. No. 8.8e-82;
                                                                                                                                                                                                                                                                                                                                                                                16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Yao N;
                                                New hepatitis C virus covalent complexes
                                                                              Claim 6; Page 76-77; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS4A-NS3 complex SEQ ID NO:16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24945 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 85.18
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP
                  WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus Synthetic.
                                                                                                                                                                                                                                                                                                                    216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BA,
                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24945;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Malcolm
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY24945
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                              ö
```

```
82 CVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                      The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS1 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the ATPase activity of NS3 The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                            5 GSVVIVGRINLSGD - - - TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                22 GSVVIVGRIILSGSSITAYSQQTRGLLGCKITSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                      the non-covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                                                                3,
                                                                                                                                                                                                                                                                                                                              86.0%; Score 874.5; DB 20; Length 665; 84.7%; Pred. No. 4.2e-81; Live 16; Mismatches 11; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yao N;
                               New hepatitis C virus covalent complexes
                                                            Page 95-97; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24946 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0094331,
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-385385/32
WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                             Local Similarity
nes 166; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                  665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24946;
                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                        62
                                                             Claim 6;
                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                 Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
```

```
NSAA PETITION OF THE PROPERTY 
                                                                                                                                                                                                                                                                                                                                present invention describes a covalent hepatitis C virus (HCV)
                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                        Claim 6; Page 97-99; 211pp; English.
```

665 AA; Sequence

5 GSVVIVGRINLSGD----TAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSTAIQIFLAI 61 22 GSVVIVGRIILSGSGSIIAXSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81 3; Gaps Query Match
86.0%; Score 874.5; DB 20; Length 665;
Best Local Similarity 84.7%; Pred. No. 4.2e-81;
Matches 166; Conservative 16; Mismatches 11; Indels 3; a ò

62 SINGVLWTVYHGAGTRTIASPKGPVT@MYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121

182

qq δ

q ö

ò

VDFIPVESLETTMRSP 197 |||:|||:|||||||||| VDFVPVESMETTMRSP 217

Search completed: August 30, 2003, 19:12:24 Job time : 45.6227 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
```

COPY 119:11 (c) 10:01 tool Compage and	OM protein - protein search, using sw model	August 30, 2003, 19:02:22 ; Search time 16.2134 Seconds (without alignments) 1168.492 Million cell updates/sec
	OM protein	Run on:

US-09-965-594-18 1017 1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197 Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 Total number of hits satisfying chosen parameters:

283308 segs, 96168682 residues

Searched:

seq length: 0
seq length: 200000000 Minimum DB Maximum DB

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
1: pir4:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

genome polyprotein polyprotein - dour polyprotein - marm nitrate/nitrite se probable periplasm probable periplasm hypothetical prote Apporterior protein hypothetical protein prote genome polyprotein probable oxidoredu genome polyprotein genome polyprotein Description SUMMARIES GNWVC3 S40770 GNWVCH GNWVTC GNWVTC GNWVTC GNWVTC GNWVTC GNWVTC GNWVTC TC6820 JC5620 JC5620 JC5620 JC5620 JC5620 JC5620 GNWVCH FRIDA H83144 B71284 AB1775 I39383 D87603 E75392 F81196 G83612 A75335 Query Match Length DB 3014 3033 3033 3033 3005 290 590 290 398 1334 452 716 377 477 477 6854. 3 68868. 3 68868. 3 68868. 3 68868. 3 68868. 3 6744. 3 6746. 3 676. 3 6868. 3 68 Score 80 80 80 80 80 Result No.

hypothetical prote	transferrin precur	irregular chiasm C	hypothetical prote	transferrin-bindin	genome polyprotein	probable exported	ferredoxin-nitrite	transferrin-bindin	surface-associated	polyketide synthas	proteinase (EC 3.4	mosA protein - Rhi	htrA-like serine p	hypothetical prote	hypothetical prote
257895	S33761	A49448	T04533	JN0821	GNWVTB	AH1030	S16603	C81832	S61441	T17410	PRLJHD	B53308	н97199	G90654	G85505
~	N	N	~	~	~	N	-	~	7	~	~	Н	~	7	~1
361	106	764	846	911	3412	492	594	910	2508	3739	323	333	433	452	452
7.7	7.7	7.7	7.7	7.7	7.7	7.6	7.6	7.6	7.6	7.6	7.5	7.5	7.5	7.5	7.5
78.5	78.5	78.5	78.5	78	78	77.5	7.7	. 77	77	77	76.5	76.5	76.5	76.5	76.5
90	31	32	33	34	35	36	37	38	39	4 0	41	42	43	44	45

ALIGNMENTS

53

---LSGDTAYAQOTRGEBGCQETSOTGRDKNOVEGEVQIVST

3 KKGSVVIVGRIN---

ò

```
A; Molecule type: genomic RNA
A; Residues: 1-3011 <INC>
A; Residues: 1-3011 <INC>
A; Cross-references: GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329738
A; Cross-references: GB:M67463; NID:g329737; PIDN:AAA45534.1; PID:g329738
B; Inchauspe, G: Zebedee, S:: Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A; Title: Genomic structure of the human proctotype strain H of hepatitis C virus: comp.
A; Reference number: A41546; MUID:92052256; PMID:1658800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F;1616-1862/Próduct: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <NNB>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4a; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A, Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: papatitis C virus genome polyprotein
F:115-Product: capsid protein M *status predicted <NSI>
F:130-729/Product: nonstructural protein NSI *status predicted <NSI>
F:130-1207/Product: nonstructural protein NSI *status predicted <NSI>
F:120-1217/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
                                                                                                                                        genome polyprotein - hepatitis C virus (strain H)
N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Description: Genomic structure of the human prototype strain H of hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1005 RRGQEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A36814; A461546
R:Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M. submitted to GenBank, July 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEEGCQETSQIGRDKNQVEGEVQIVSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          C, Accession: A40244
R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.4%; Score 848.5; DB 1;
81.9%; Pred. No. 9.7e-69;
tive 10; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 81.98
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Species: hepatitis C virus
                                                                                                                                                                                                                                               C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Reference number: A36814
A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A. Experimental source: isolate HG-JI

C. Superfamily: hepatitis C virus genome polyprotein

C. Superfamily: hepatitis C virus genome polyprotein

C. Superfamily: hepatitis C virus genome

C. Superfamily: hepatitis C virus genome

F. 2-115/Product: capsid protein C *status predicted <CPC>
F. 116-191/Product: capsid protein M *status predicted <EPM>
F. 192-396/Product: major envelope protein E *status predicted <MEE>
F. 390-729/Product: nonstructural protein NS2 *status predicted <NS2>
F. 730-1006/Product: nonstructural protein NS2 *status predicted <NS2>
F. 1007-1615/Product: nonstructural protein MS2 *status predicted <NS2>
F. 1230-1231/Region: nucleotide-binding motif A (P-loop)

F. 1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Cross-references: EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221587
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The S'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1005 RKGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  173
                                                                                                                                                                                                                                               1005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                         54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: $40770; PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                              A;Cross-references: GB:D00831; NID:9221511; PIDN:BAA00705.1; PID:9221512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 853.5; DB 1
Pred. No. 3.4e-69;
8; Mismatches 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             R;Okamoto, H. submitted to the EMBL Data Library, March 1992 A;Reference number: $40770 A;Accession: $40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1185 CIRGVAKAVDFIPVESLETIMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                           STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83.9%;
82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: genomic RNA A; Residues: 1-513 < OK2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA
A; Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity
Matches 169; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          174
                                                                                                                                                                                                                                                                                                                                                      174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                        g
                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                        g
   유
                                                                         ò
```

113

Length 3011; Indels

```
F;1616-1862/Product:
F;1863-2013/Product:
F;2014-3010/Product:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A Experimental Source: HCV-JT,
A:Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
A:Note: sequence extracted from NCBI backbone (NCBIN:106206, NCBIP:106207)
C:Superimanly: hepatilis C virus genome polyprotein
E:116-191/Product: capsid protein C *status predicted <PRD>
F:105-191/Product: envelope protein M *status predicted <PRD>
F:106-129/Product: nonstructural protein NS2 *status predicted <NS2>
F:1007-1615/Product: nonstructural protein NS2 *status predicted <NS2>
F:1007-1615/Product: nonstructural protein NS2 *status predicted <NS2>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                C. Reywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura F;1-115/Product: capsid protein C *status predicted <CPC>
F;16-191/Product: envelope protein M *status predicted <CPC>
F;192-189/Product: major nevelope protein M *status predicted <MFE>
F;390-729/Product: major nevelope protein NSI *status predicted <NSI>
F;300-1006/Product: nonstructural protein NSI *status predicted <NSI>
F;1007-1615/Product: nonstructural protein NSI *status predicted <NSI>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (strain JT)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructup protein NS4s; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: A45573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,214,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A:Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     R; Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Virus Res. 23, 39-53, 1992
A; Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: A; Reference number: A45573; MUD: 92295714; PMID:1318627
A; Reference number: A45573
A; Ratus: preliminary
A; Molecule type: DNA
A; Residues: 1-3010 ATM>
A; Cross-references: GB: D11168; GB: D01171; NID: 9221612; PIDN: BAA01943.1; PID: 9221613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3 KKGSVVIVGRIN------LSGDTAYAQOTRGEEGCQETSQTGRDVVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82.4%; Score 837.5; DB 178.4%; Pred. No. 9.8e-68;
                                                                                                                                                                        A;Cross-references: GB:M84754
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 82.4
Best Local Similarity 78.4
Matches 160; Conservative
                                                                                                     A; Molecule type: genomic RNA
A; Residues: 1-3010 <CHE>
                                                                         A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
```

```
A; Molecule type: genomic RNA
A; Residues: 1-3010 A; Rosidues: 1-3010 A; Rosidues: 1-3010 A; Rosidues: 1-3010 A; Residues: 1-3010 A; Residues: 1-3010 CTAX>
A; Residues: 1-3010 A; Residues: 1-3010 A; Residues: 1-3010 A; Cross-references: EMBL:M58335; NID:g329770; PIDN:AA72945.1; PID:g329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstruct
F; 2-115, Product: envelope protein E *status predicted <ERD>
F; 192-389, Product: major envelope protein NS1 *status predicted <NES>
F; 1007-1615, Product: nonstructural protein NS1 *status predicted <NS2>
F; 1007-1615, Product: nonstructural protein NS2 *status predicted <NS2>
F; 1107-1117, Region: nucleotide-binding motif B (P-loop)
F; 1110-1117, Region: DEXH motif
F; 1116-11319, Region: DEXH motif
F; 1166-1862, Product: nonstructural protein NS4* *status predicted <NSA>
F; 2014-3010, Product: nonstructural protein NS4* *status predicted <NSS>
F; 1204-3010, Product: nonstructural protein NS4* *status predicted <NSS>
F; 1206-209, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1257, 2011, 207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus genome polyprotein (EC 3.4.21.98) (nonstr N; Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr protein Ns4a; nonstructural protein Ns4b; nonstructural protein Ns5 C; Species: hepatitis C virus C; orders: 1-mar-1992 *sequence_revision 31-mar-1992 *text_change 19-Jan-2001 C; Accession: A38465 R; Takamizawa, A.; Mori, C.; Puke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, J., Virol. 65, 1105-1113, 1991 A; Takamizawa, A.; Mori, C.; Puke, I.; Hanabe, S.; Murakami, S.; Fujita, J.; Onishi, J.; Virule: Structure and organization of the hepatitis C virus genome isolated from hu A; Reference number: A38465; MuID:91140698; PMID:1847440
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RRGKEILLGPADSLEGRGLRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATOSFLATCVNGVCWTVYHGAGSKTLAAPKGPITOMYTNVDQDLVGWPKPPGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1005 RRGREILLGPADSIEGGGWRLLAPITAYAQOTRGLLGCIVTSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CGSSDLYLVTRHADV I PVRRRGDSRGSLLSPRP I SYLKGSSGGPLLCPAGHAVGI FRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIN------LSGDTAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN------LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
nonstructural protein NS4a *status predicted <N4A> nonstructural protein NS4b *status predicted <N4B> nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ٠.
م
                                                                                                                                                                                                                                                          6
                                                                                                                                                                        Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18; Indels
                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                    Score 827.5; DB 1;
Pred. No. 8e-67;
20; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81.0%; Score 823.5; DB 1
76.5%; Pred. No. 1.9e-66;
Live 21; Mismatches 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                             81.48;
                                                                                                                                                                                                                   77.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Siminarity (0.5)
Matches 156; Conservative
                                                                                                                                                                        Query Match
Best Local Similarity 77.0
Matches 157; Conservative
```

ô 셤

```
C. Steywords: ATP: 91ycoprotein; hydrolass; nucleotide binding; P-loop; polyprotein; se Fig-115/Product: capsid protein ( #status predicted <CPC>
Fig-115/Product: capsid protein ( #status predicted <CPC>
Fig-115/Product: envelope protein M *status predicted <MED>
Fig-2-389/Product: envelope protein E *status predicted <MED>
Fig-2-389/Product: major envelope protein NSI *status predicted <NSI>
Fig-100/Froduct: nonstructural protein motein M (P-loop)
Fig-100/Froduct: nonstructural protein motein M (P-loop)
Fig-100/Froduct: nonstructural motein motif B (P-loop)
C;Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 23-Mar-2001 C;Accession: S18030, S33570; A48325; S18029 K; Murakami, S. S. Sheko, S.; Masashi, U.; Kobayashi, K.; Murakami, S. Submitted to the EMBL Data Library, September 1991 A;Bescription: A whole genome of hepatitis C virus cDNA was isolated from a single pa A;Reference number: S18028 A;Reference number: S18038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A.Experimental source: isolate JRI from an individual
R.Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A.TILE: Sequence analysis of putative structural regions of hepatitis C virus isolat
A.R.Reference number: A48332; MUID:93119270; PMID:8380322
A.Reference number: A48332; MUID:93119270; PMID:8380322
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome polyprotein - hepatitis C virus (isolate EUH1480)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru protein N84a; nonstructural protein N84b; nonstructural protein N85 C; Species: hepatitis C virus
C;Species: hepatitis C virus
C;Species: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Richamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A. Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomina
A.; Reference number: JC5620; MUID: 97366593; PMID: 9222423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N8A>
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5
F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Molecule type: genomic RNA
A; Residues: 1-547, T', 549-621, V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782
A; Cross-references: EMBL:X61591
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Note: this sequence is inconsistent with the nucleotide translation A; Note: the authors translated the codon AGG for residue 43 as Pro, TGG for As Translated for TS Ser As Translated for TS Ser A; Note: sequence extracted from NCBI Dackbone (NCBIN:121747, NCBIP:121748) C; Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1065 ATQSFLATCVNGVCWTVYHGAGSKTLAGPKGPINQMYTNVDQDLVGWQAPSGAASLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: genomic RNA
A;Residues: 1-3010 <HON>
A;Cross-references: EMBL:X61596; NID:g59478; PIDN:CAA43793.1; PID:g59479
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79.9%; Score 812.5; DB 1; ilarity 76.0%; Pred. No. 1.9e-65; Conservative 20; Mismatches 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity
Matches 155; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Accession: JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1005
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: genomic RNA
A; Residues: 2650-2707 (KA2>
A; Residues: 2650-2707 (KA2>
A; Residues: 2650-2707 (KA2>
A; Experimental source: Japanese isolate
C; Comment: The cleavage sites of this polyprotein have not been determined.
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis c virus predicted <PED>
C; Superfamily: hepatitis protein E * Status predicted <PED>
C; Superfamily: hepatitis detains motif B
C; Superfamily: hepatitis motif B
C; Superfamily: hepatitis detains d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 8
$18030
genome polyprotein - hepatitis C virus (isolate JK1)
N;Contains: capsid protein C; envelope protein N; hepaclvirin (EC 3.4.21.98) (nonstructu
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4s, nonstructural protein NS5
D;Dotein NS4s, nonstructural protein NS4b; nonstructural protein NS5
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C;Accesalon: A3923; 80006
C;Accesalon: A3923; 80006
C;Accesalon: A3923; 80006
C;Accesalon: A39253; MUID:91088550; PMID:2175903
A;Reference number: A39253; MUID:91088550; PMID:2175903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R;Kato, N.; Obkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A;Fithe: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variancesion: PS0086
A;Reference number: PS0085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N85>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Molecule type: genomic RNA
A;Residues: 1-3010 <KAI>
A;Cross_references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81.0%; Score 823.5; DB 1; 75.5%; Pred. No. 1.9e-66; tive 23; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CTRGVAKAVDFIPVESMETTMRSP 1208
174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   larity 75.5%;
Conservative 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ouery Match
Best Local Similarity
Matches 154; Conserv
```

54

ò g ò g ö 셤 ö q S

```
Across references: GB:D10562; GB:D90518; NID:g221523; PIDN:BAA01418.1; PID:g221524
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein; plycorotein; hydrolase; nonstruc
C:Superfamily: hepatitis C virus genome polyprotein; plycoduct: capsid protein C *status predicted <CPC>
E:116-191/Product: envelope protein E *status predicted <NES>
E:1301-2189/Product: nonstructural protein NSI *status predicted <NSI>
E:1011-1619/Product: hepatitiin *status predicted <NSI>
E:1011-1619/Product: hepatitis <NSI *status *status
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C;Species: hepatitis C virus
C;Date: 19-Way-2000 #sequence_revision 19-May-2000 #text_change 17-Nov-2000
C;Date: 19-Way-2000 #sequence_revision 19-May-2000 #text_change 17-Nov-2000
C;Accession: J01303
B;Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurat, K.; Iizuka, H.; Machida, A.; Miyakawa J. Gen. Virol. 72, 2697-2704, 1991
A;Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a A;Reference number: J01303; MUID:92044440; PMID:1658196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A. Experimental source: 1800944; NID:9221650; PIDN:BAA00792.1; PID:9221651
A. Experimental source: 1801ate HC-J6 from a Japanese individual
C. Superfamily: hepartitis C virus genome polyprotein
C. Superfamily: hepartitis C virus genome polyprotein: serine proteinses; tr
E; 2-115/Product: capsid protein C *status predicted <CPC>
F; 116/19/Product: envelope protein M *status predicted <CPC>
F; 116-19/Product: major envelope protein E *status predicted <MED>
F; 130-733/Product: monstructural protein NS1 *status predicted <NS1>
F; 134-1010/Product: nonstructural protein NS2 *status predicted <NS2>
F; 101-1619/Product: hepacivirin *status predicted <NS3>
F; 1316-1321/Region: nucleotide-binding motif B
F; 1320-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    F;1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <N0S>
F;2018-3033/Product: nonstructural protein NS5 #status predicted <N0S>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F:1620-1866/Product: nonstructural protein NS4a #status predicted <N4A>
F:1867-2017/Product: nonstructural protein NS4b #status predicted <N4B>
F:2018-3013/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,233,299,305,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         66.4%; Score 675; DB 1; Length 3033;
69.8%; Pred. No. 6.8e-53;
tive 24; Mismatches 30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66.3%; Score 674; DB 1; Length 3033; 68.7%; Pred. No. 8.4e-53;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       27; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity 69.8%
Matches 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 68.74
Matches 123; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     F;1320-1323/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA A; Residues: 1-3033 < OKA>
                                                               ₹¥
A; Molecule type: mRNA
A; Residues: 2678-2729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              J01303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                genome polyprotein - hepatitis C virus (strain HC-JB)
N:Contains: capsid protein C: envelope protein M: hepacivirin (EC 3.4.21.9B) (nonstructur protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 (Species: hepatitis C virus (C:Species: hepatitis C virus (C:Species: hepatitis C virus (C:Accession: A40250; po0397; po0559
R:Okamoto, H.; Kurai, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda, S.; Airology 18B, 331.341, 1992
A:Title: Full-length sequence of a hepatitis C virus genome having poor homology to repc A:Reference number: A40250; MUID:92230232; PMID:1314459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Experimental source: isolate E-b12
R; Kato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno Biochem. Blophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural 181, 279-285, papan.
A; Reference number: P00554; MUID: 92068204; PMID: 1720309
                                                                                                                                    A;Experimental source: genotype 5a, which predominates in South Africa A;Note: the translation of the nucleotide sequence is not complete in this paper A;Note: the translation of the nucleotide sequence is not complete in this paper (S.Superfamilly: hepatilis C virus genome polyprotein (S.Superfamilly: hepatilis) appropriet (S.Superfamilly: hepatilis predicted CPC) (S.Superfamilly: hepatilis protein (S.Superfamilly: hepatilis (S.Superfamill): S.Superfamilly: hepatilis (S.Superfamill): S.Superfamilly: hepatilis (S.Superfamill): S.Superfamilly: S.Superfamills: S.Super
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA
A; Residues: 1-3033 <00X2>
A; Residues: 1-3033 <00X2>
A; Casidues: 1-3033 <00X2>
A; Casidues: 1-3033 <00X2>
A; Casidues: 1-3033 <00X2>
A; Casidues: 1-3033 <00X2
B; Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P. J. Gen. Virol. 73, 113-1144, 1992
A; Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to A; Reference number: PQ0393; MUID:92268871; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1006 RRGREIFLGPADDIKTSGWRLLAPITAYAQQTRGVLGAIVLSLTGRDKNEAEGEVQFLST 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1126 CGSADLYLVTRHADVIPARRRGDTRASLLSPRPISYLKGSSGGPIMCPSGHVVGVFRAAV 1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           114 CGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 ATQIFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F.1617-1863/Product: nonstructural protein NS4a *status predicted <NAA> F.18617-2014/Product: nonstructural protein N854 *status predicted <N4B> F.2015-3014/Product: nonstructural protein N85 *status predicted <NS5> F;2210-2249/Region: interferon sensitivity determining *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 3014;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               73.1%; Score 743.5; DB 1; 68.6%; Pred. No. 3.8e-59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1186 CTRGVAKALEFVPVENLETTMRSP 1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA
A; Residues: 2678-2754 <CHA>
A; Cross references: DDBJ:D10134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 68.6%
Matches 140; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1317-1320/Region: DEXH motif
                                              A; Residues: 1-3014 <CHA>
A; Cross-references: GB:Y13184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: PQ0397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
```

ö

78

ö

Gaps

9

Oy 53 TATOTFLATSINGVLMTVYHGAGTRTIASPKGPVTOMYTNVDKDLVGWQAPQGSRSLTPC 112 Db 1006 TSTTRSMGTCVNGVMTTTYHGSNARTLAAQMGPVNSRWMSASDDVAYPLPVGARCLEPC 1065 Qy 113 TCGSSDLXLVTRHADVIPVRRGDSRGSLLSPRPISYLRGSSGGPLLCP 161 Db 1066 KCQPQGVWVIRND-GALCHGTLGRTVELDLPAPELCDFRGSSGSPLLCD 1112 Qy 162 AGHAVGTRAVSTRGVARAVDFIPVESLETTWRSP 197	RESULT 14 B81104 Intrice sensor protein (EC 2.7.3) NWB1249 [similarity] - Neisseria meningit C;Species: Neisseria meningitidis C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001 C;Species: Neisseria meningitidis C;Date: 31-Mar-2000 C;Date: 31-Mar-2000 C;Species: Neisseria meningitidis C;Date: 31-Mar-2000 R;Tettelin, H.; Vandhers, N.J.; Heldelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. Ti, H.; Oin, H.; Vandheran, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000 A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A;Rocession: B81104 A;Status: preliminary A;Molecule type: DNA A;Rocession: 1-500, Approx.	erences tal sou 1249 ly: nit autoph	Ouery Match Ouery Match 9.18; Score 92.5; DB 2; Length 590; Best Local Similarity 21.38; Pred. No. 1.5; Matches 46; Conservative 26; Mismatches 79; Indels 65; Gaps 6; Qy 28 EEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRTIASFRGPVT 87	QY B8 QMYTHVDXDLVGWQAPQ	RESULT 15 CB1911 nitrate/nitrite sensor protein (EC 2.7.3) NWA1418 [similarity] - Neisseria meningit C;Species: Neisseria meningitidis C;Species: Neisseria meningitidis C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001 C;Accession: CB1911 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre Nature 404, 502-506, 2000 A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis 22491 A;Reference number: AB1775; MUID:20222556; PMID:10761919 A;Status: preliminary
Qy 19 TAYAOOTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78	RESULT 12 T08841 polyprotein - douroucouli hepatitis GB virus A CiSpecies douroucouli hepatitis GB virus A CiSpecies; douroucouli hepatitis GB virus A CiSpecies; douroucouli hepatitis GB virus A CiDate: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000 C;Accession: T08841 SErker, JC:; Desai, S.H.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J. Gen. virol. 79, 41-45, 1998 A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A;Reference number: 216486; MUID:99120818; PMID:9460920 A;Accession: T08841 A;Status: translated from GB/EMBL/DDBJ A;Molecule type: mRNA A;Molecule type: mRNA A;Molecule type: mRNA A;Residues: 1-3005 CERA A;Cross-references: EMBL:AF022425; NID:92828599; PIDN:AAC40502.1; PID:92828600 C;Superfamily: hepatitis C virus genome polyprotein C;Reywords: polyprotein	Ouery Match 24.5%; Score 249; DB 2; Length 3005; Best Local Similarity 33.5%; Pred. No. 5.4e-14; Matches 55; Conservative 30; Mismatches 69; Indels 10; Gaps 3; Qy 33 ETSQTGRDKNOVEGEVOIVSTATQTFLATSINGVLWIVYHGAGTRIIASPRGPVTQWTN 92 : : :: : : : : : :: :	QY 93 VDKDLVGWQAPQGSRSLIPCTCGSSDLXLVTRHADVIPVRRGDSRGSLLSPRPISYLKG 152	TESULT 13 TOB839 TOB839 TOB920 A TESULT 1399 TOB920 A TESULT 1399 TOB920 A TESULT 1999 TOB920 A TESULT 1999 TOB920 A TESULT 1999 TOB920 A TESULT 1999 TOB930 TOB839 TOB930 TOB839 TOB930 TOB930 TOB930 TOB930 TO	Status: tre Molecule Ly Residues: 1 Cross-refer Cross-refer Superfamily Keywords: Local Matches Matches Matches Matches

```
A: Molecule type: DNA
A; Residues: 1-590 <PAR>
A; Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84658.1; PID:g738007
A; Experimental source: serogroup A, strain 22491
C; Genetics: NAA1418
C; Superfamily: nitrate/nitrite sensor protein narX
C; Superfamily: nitrate/nitrite sensor protein narX
C; Seywords: autophosphorylation; phosphoprotein; phosphotransferase; gF; 395/Active site: His (phosphohistidine intermediate) *status predicted
                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                               ....-VIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTR--- 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 319 DLGKYHEEIFPIEYQNETLGRLLLSFPNGISLDEDDRILLQTLGRQLGVSLAGAKQEEEK 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     88 QMYTNVDKDLVGWQAPQ-----127
                                                                                                                                                                                                                                                                                                                                                                                           28 EEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRTIASPKGPVT 87
                                                                                                                                                                                                                                                                                 Query Match 9.1%; Score 92.5; DB 2; Length 590;
Best Local Similarity 21.3%; Pred. No. 1.5;
Matches 46; Conservative 26; Mismatches 79; Indels 65; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          gg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
```

Search completed: August 30, 2003, 19:20:30 Job time : 17.2134 secs

oy O

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 9.75674 Seconds (without alignments) 949.524 Million cell updates/sec

Run on:

US-09-965-594-18 1017 1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	P26664 h genome po	c	h genome		h genome	h genome	h genome					Q13685 homo sapien	P29837 1 genome po									Q01299 t genome po		Q07607 rhizobium m	P23931 escherichia	P03314 y genome po	P19901 y genome po	~		~		qqib 1	P11033 mus musculu
SUMMARIES	ai ai	POLG_HCV1	POLG_HCVH	POLG_HCVTW	POLG_HCVJT	POLG_HCVBK	POLG_HCVJA	POLG_HCVJ8	POLG_HCVJ6	HHOA_ARATH	PAAD_PSEAE	DEG1_ARATH	AAMP_HUMAN	POLG_LANVT	POLG_TBEVW	TRFE_HORSE	ICCR_DROME	TB11_NEIMB	POLG_TBEVS	GRAK_MOUSE	NIR_SPIOL	PTPO_MOUSE	POLG_TBEVH	VPRT_SMRVH	MOSA_RHIME	MLTD_ECOLI	POLG_YEFV1	POLG_YEFV2	Y136_TREPA	WDR9_HUMAN	NX1B_BOVIN	VP18_HUMAN		GRAD_MOUSE
	DB	-	-	Н	7	-	~	-	٦	-	-	-	-	ч	Н	Н	Н	-	-	~	~	H	-	-	-	Н	-	-	-	-	-	-	-	-
	Length	3011	3011	3010	3010	3010	3010	3033	3033	321	209	437	452	3414	3414	206	764	911	3412	263	294	1705	3414	323	333	452	3411	3411	485	2269	467	973	1165	248
•	Query Match	84.0	83.4	82.4	81.4	81.0	81.0	9 .99	66.3	9.8	٠	•	•	7.8	٠	•	•	•	•	•	7.6	•	•	7.5	•	7.5	7.5	•	7.4		7.4	7.4	7.4	7.3
	Score	854.5	848.5	837.5	827.5	823.5	823.5	675	674	87	85.5	83.5	83	79.5	79	78.5	78.5	78	78	77.5	77	77	77	76.5	76.5	76.5	16	16	δ.	75.5		75	7	74.5
	Result No.	1	7	٣	4	5	9	7	æ	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	36	27	28	29	30	31	32	33

P26596 lactococcus P97608 rattus norv P54748 rattus norv P5486 ustilaso ma Q96qs1 homo sapien Q88621 mus musculu P08169 bos tarus P28863 oryctolagus Q28146 bos taruus Q04538 t genome po Q67531 aquifex aeo	•
MCP_BPF41 OPLA_RAT OPLA_RAT OPLA_RAT DPM1_USTWA PHWX_HUMAN FXH1_MOUSE WRT_BOVIN MAG03_RABIT NX.1A_BOVIN NX.1A_BOVIN FLII_AQUAE FLII_AQUAE POOLS POWYL	
пенапанана	•
301 1288 8448 294 320 401 2499 1530 3415 861	•
	:
74.5 74.5 73.5 73.5 73.5 72.7 73.5 73.5 73.5 73.5 73.5 73.5 73.5 73) !
@@@@@@################################	

ALIGNMENTS

RESULT 1 POLG_HCV1 ID POLG_HCV1 STANDARD; PRT; 3011 AA.	Pibbo4; 01-AuG-1992 (Rel. 23, Created) 01-AuG-1992 (Rel. 23, Last sequence update)	13 SEF 1003 (REL. 4, Lact annotation update (Core protein) (P22); Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2	(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22); Protease/helicase NS3 (P70) (Hepacivirin)	(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein	NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)]. Hepatitis C virus (isolate 1) (HCV).	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;	ncpart trus.	[1] SEOUENCE FROM N.A.	MEDLINE-91172826; PubMed-1848704;	Choo U. L., Alduman N.H., Han J.H., Berger N., Lee C., Dong C., Gallegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,	Bradley D.W., Kuo G., Houghton M.; "Constin or wires".	Proc. Natl. Acad. Sci. U.S.A. 88:2451-2455(1991).	-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4B AND NS4B ARE HYDRODOURIC SHICKFORTHS A POSSTRIP MEMBRANE-RELATED FUNCTION	NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.	 -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Ago or Glu in the P6 	position, Cys or Thr in Pland Serior That in Pl'.		-!- SUBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A	PROTEIN WAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF	PROTEIN C AND MANA. -:- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.	This currence and a consequent of it a second through a collaboration	smiss from energy is copyright. It is produced through a correction the Swiss Institute of Bioinformatics and the EMBL outstation	the European Bioinformatics Institute. There are no restrictions on its	modified and this statement is not removed. Usage by and for commercial	entitles requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license(lsb-sib.ch).		EMBL; M6231; AAA456/6.1; PIN: A39166; GNWAC3.	PUB; IAIV; IO-FED-33. PUB; IHEI; 25-NOV-98.	, G	InterPro; IPR002522; HCV_capsid.
RESULT POLG_H	SEP E	0 E	DE DE	DE	OS	88	38	X Y	X a	\$ \$	8 5	2 2	ខ្លួ	ខ្ល	ပ္ပင္	ខ្លួ	ខ	ខ្ល	388	ខ្លួ	ខ្ល	ខ	95	30	ဗ္ဗဗ	ပ္ပ	22.2	2 2 2	5 5 5	ž Ž

```
PRT; 3011 AA
                                                                                                                                                                                                                                 174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Structure 6:89-100(1998).
                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus.
NCBI_TaxID-11108;
                                                                                                                                                                                                                                                                                                                                                POLG_HCVH
                                                                                                                                                                                                                                                                                                        RESULT 2
POLG_HCVH
                                                                                                                                                                 g
                                                                                            8
                                                                                                                                     ò
                                                                                                                                                                                                             ò
                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CELLOLAR AMINOPEPTIDASE.
CAFSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/R2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/R2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4/R4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4/R4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4/R4 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ( POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ( POTENTIAL )
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RELAY SYSTEM (BY
RELAY SYSTEM (BY
RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC. . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       65F8C9447FCE5AF9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 854.5; DB 1 Pred. No. 2.1e-71;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GICNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED (GLCNAC N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LINKED
                                                                                        InterPro; IPR00145; HCV_NS4a.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001566; HCV_RGRP.
InterPro; IPR001506; Helicase_C.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR0071094; RNA_pol_PSvir.
Pfam; PF01539; HCV_core; I.
Pfam; PF01539; HCV_core; I.
Pfam; PF01539; HCV_NS1; I.
Pfam; PF01539; HCV_NS1; I.
Pfam; PF01539; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE
                                                                                                                                                                                                         InterPro; IPR007094; RNA_pol_PS-
Pfam; PP01543; HCV_capsid; I.
Pfam; PP01542; HCV_capsid; I.
Pfam; PP01539; HCV_env; I.
Pfam; PP01539; HCV_env; I.
Pfam; PP01539; HCV_MS1; I.
Pfam; PP01005; HCV_MS3; I.
Pfam; PP01006; HCV_MS3; I.
Pfam; PP01006; HCV_MS4s; I.
Pfam; PP00201; HCV_MS4s; I.
Pfam; PP00201; HCV_MS4s; I.
Pfam; PP00271; helicase_C; I.
Pfam; PP0099; Viral_RARP; I.
ProDom; PP0186062; HCV_MS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         84.0%;
nilarity 82.8%;
Conservative 9
                                                      HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA; 327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3011
369
1083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1107
11165
11237
1319
196
209
234
305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    645
2041
2077
                                                                                                                                                                                                                                                                                                                                                                                                                                     ProDom; PD186062; HCV_r
SMART; SM00487; DEXDC;
                                                                            IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3D-structure.
INIT_MET
                                                      InterPro;
InterPro;
                    interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local S
Matches 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
```

3 KKGSVVIVGRIN------LSGDTAYAQOTRGEEGCQETSQTGRDKNQVEGEVQIVST

Gaps

6

Indels

9; Mismatches

169;

ö

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                    01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 42, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P72); Nonstructural protein NS58 (P56); Nonstructural protein
NS58 (P60) (P70) (RNA-directed RNA Polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-98154321; PubMed-9493270; Kim J.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A., Murcko M.A., Lin C., Caron P.R.; "Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              H.V., Weber P.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -:- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-:- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CATALYTIC ACTIVITY: Hydrolygis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

CATALYTIC ACTIVITY: Nucleoside triphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.
-!- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
-!- FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-!- FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENTIAL ROLE IN THE VIRUS REPLICATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Prince A.M.; "Genomic strain H of hepatitis "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hoong Z., Kwong A.D., Le H.V
"Structure of the hepatitis C virus RNA helicase domain.
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA) (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ---
```

```
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 848.5; DB 1;
Pred, No. 7.5e-71;
                 GLCNAC.
GLCNAC.
GLCNAC.
GLCNAC.
GLCNAC.
GLCNAC.
GLCNAC.
GLCNAC.
               N-LINKED
       LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              300

4444

4410

5376

5376

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6536

6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1382
1389
1397
                                           CARBOHYD
TURN
HELIX
TURN
HELIX
TURN
HELIX
TURN
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
TURN
TURN
TURN
TURN
TURN
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
TURN
TURN
TURN
TURN
TURN
STRAND
STRAND
STRAND
STRAND
STRAND
STRAND
TURN
TURN
TURN
TURN
STRAND
TURN
STRAND
STRAND
STRAND
STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                 CARBOHYD
                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HELIX
HELIX
STRAND
HELIX
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TURN
HELIX
TURN
HELIX
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                     InterPro; 1PR002521; HCV_core.

R InterPro; 1PR002521; HCV_core.

R InterPro; 1PR002521; HCV_core.

R InterPro; 1PR002519; HCV_core.

R InterPro; 1PR002519; HCV_Lenv.

R InterPro; 1PR002519; HCV_NS1.

R InterPro; 1PR001409; HCV_NS3.

R InterPro; 1PR001409; HCV_NS4a.

R InterPro; 1PR001409; HCV_NS5a.

R InterPro; 1PR00105166; HCV_RARA.

R InterPro; 1PR001056; HRA_pol_DS_PS.

R InterPro; 1PR001056; HRA_pol_DS_PS.

R InterPro; 1PR001056; HCV_RARA.

R InterPro; 1PR001056; HCV_RARA.

R InterPro; 1PR001056; HCV_NS5a.

R Pfam; PF01506; HCV_NS2; 1.

R Pfam; PF01506; HCV_NS3; 1.

R Pfam; PF01006; HCV_NS4a; 1.

R Pfam; PF01006; HCV_NS4a; 1.

R Pfam; PF01006; HCV_NS4a; 1.

R Pfam; PF001006; HCV_NS4a; 1.

R Pfam; PF001006; HCV_NS5a; 1.

R Pfam; PF0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROFEDITY C PROCESSING BY THE HOST SIGNAL PEPTIDASES.
SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39.
SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC. ..) (POTENTIAL)
N-LINKED (GLCNAC. ..) (POTENTIAL)
N-LINKED (GLCNAC. ..) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REMOVED FROM CAPSID PROTEIN C
CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEASE/HELICASE NS3
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5B.
POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMII
CHARGE RELAY SYSTEM (BY SIMII
ATP (POTENTIAL).
CHARGE RELAY SYSTEM (BY SIMII
ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
                                                                                                                                                                  EMBL, M67463; AAA45534.1; PTR, AA5814; GNWVCH. PDB; 1HEI; 25-NOV-98. PDB; 1A1V; 16-FEB-99. PDB; 1A1V; 17-TUN-98. MEROPS; S29.001; TRANSFAC; T04155; InterPro; IPR001410; DEAD. InterPro; IPR002521; HCV_caps! InterPro; IPR002521; HCV_caps! InterPro; IPR002531; HCV_caps! InterPro; IPR002531; HCV_caps! InterPro; IPR002531; HCV_NS1. RICEPPro; IPR002531; HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            192
384
747
810
1105
1105
1107
1107
11165
11165
11165
1230
1316
1316
234
245
11165
11165
11165
1230
234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                      114 CGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                           54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                            01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-5PP-2003 (Rel. 42, Last annotation update)
15-5PP-2003 (Rel. 42, Last annotation update)
6enome polyprotein [Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5B (P77); Nonstructural protein NS5B (P77); Nonstructural protein NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
  9; Gaps
                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            {RNA}(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
  18; Indels
                                                                                                                                                                                                            PRT; 3010 AA
10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          send an email to license@isb-sib.ch).
                                                                                                                                                  174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL; M84754; -; NOT_ANNOTATED_CDS
Matches 167; Conservative
                                                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDB; 1N64; 25-FEB-03
PDB; 1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; A40244; GNWVTW
                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-31645;
                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                                                                                                                           POLG_HCVTW
P29846;
                                                                             ದ್ದ
                                                                                                   õ
                                                                                                                     g
                                                                                                                                        ò
                                                                                                                                                           g
                                                                                                                                                                                                            õ
                                      셤
                                                           ó
```

fnterPro; IPR001410; DEAD.

MEROPS; S29.001; -.

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                              SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Iransmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY)
SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MW; AAD267D55CDFE215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         82.4%; Score 837.5; 78.4%; Pred. No. 8e-7 tive 18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                        Interpro: IPR007095; RNN__POL_DS_PS.
Interpro: IPR007094; RNN_POL_DS_PS.
Interpro: IPR007094; RNN_POL_DS_PS.
Interpro: IPR007094; RNN_POL_DS_PS.
Interpro: IPR074; HCV_core; 1.
Pfam; PF0150; HCV_Env; 1.
Pfam; PF0150; HCV_NS; 1.
Pfam; PF01006; HCV_NS; 1.
Pfam; PF01001; HCV_NS; 1.
Pfam; PF01001; HCV_NS5; 1.
Pfam; PF01001; HCV_NS5; 1.
Pfam; PF01506; HCV_NS5; 1.
Pfam; PF00106; HCV_NS5; 1.
Pfam; PF0010998; Viral_RdRP; 1.
                                                                                                HCV_NS5a.
                                                                            HCV_NS4b.
                                                                 HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           645
2041
2240
2240
2529
2788
AA; 327047 N
                                                HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
Interpro; IPR002519; H
Interpro; IPR002511; H
Interpro; IPR002518; H
Interpro; IPR004109; H
Interpro; IPR004109; H
                                                                                                               InterPro; IPR002166;
                                                                              IPR001490;
                                                                                                InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                              3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      160;
                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
```

```
----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                  erPro; IPR007095; RNA_pol_DS_PS.
erPro; IPR007094; RNA_pol_DS_PS.
n; PF01543; HCV_capsid; 1.
pp01543; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       20;
                                                                                                   HCV_NS4b.
HCV_NS5a.
                                                                                                                                                                                                                                                                                           F01506; HCV_NS5a; 1.
FF00271; helicase_C; 1.
F00998; Viral_RdRP; 1.
PD186062; HCV_NS1; 1.
SM00487; DEXDC; 1.
                                                                                      HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      623
645
2041
2077
2240
2529
2788
AA; 326573 M
                                                                                                                                                                                                                                 HCV_NS2; 1.
HCV_NS3; 1.
HCV_NS4a; 1
HCV_NS4b; 1
                                                                                                                                                                                         HCV_core;
HCV_env; 1
HCV_NS1; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches 157; Conservative
                                                                     IPR004109; F
IPR000745; F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                               PF02907;
PF01006;
PF01001;
PF01506;
                                                                                                                                                                                                                                                                                                                          Pfam; PF00998;
            InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                              InterPro;
                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                         ProDom;
SMART; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEM outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license@labshib.ch.
                                                                                                                                1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVST 1064
                                                       ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                               CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                    01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
64-67 (Rel. 42, Last annocation update)
65-67 (Gontains: Capsid protein C (Core protein) (P22);
67-68) (GP76) (RMI); Protein F1 (GP32) (GP35); Envelope glycoprotein E2 (GP83) (GP35); Envelope glycoprotein E2 (GP83) (GP35); Envelope glycoprotein E2 (GP83) (GP3) (GP3) (GP3) (GP35); Envelope glycoprotein E2 (GP83) (GP3) 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    {RNA}(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LIPOPROTEIN ENVELOPE, THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-92295714; PubMed-1318627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; D11168; BAA01943.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIR; A45573; A45573.
PDB; 1A1Q; 25-MAR-98.
PDB; 1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROPS; U39.001;
                                                                                                                                                                                                                                                                             POLG_HCVJT
                                                                                                               114
                                                                                                                                         ద
                                                                                                                                                                                                                                                                                           셤
                                                                              q
                                                                                                                                                                       ð
                                                    ò
                                                                                                            οy
```

```
CELLULAR AMINOPEPIDASE.
CAPEID PROTEIN (POTEWITAL).
MAJOR ENVELOPE PROTEIN E (POTEWITAL).
MAJOR ENVELOPE PROTEIN E (POTEWITAL).
NONSTRUCTURAL PROTEIN NSI,F22 (POTEWITAL).
PRON STRUCTURAL PROTEIN NSI,F22 (POTEWITAL).
PROTEASE/HELICASE NS3 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
RONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
RONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
RONSTRUCTURAL PROTEIN NS48 (POTEWITAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
Polyprotein, diycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease; 3D-structure. 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                               SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 827.5; DB 1; Length 3010;
Pred. No. 6.9e-69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (
CHARGE RELAY SYSTEM (
ATP (POTENTIAL) .
DECH BOX.
N-LINKED (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDINE-9822846; PubMed-9568891;
A Steinkueler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.,
Tankueler C., Tomel L., de Francesco R., Kuo L.C., Chen Z.,
T. Complex of NS3 protease and NS4 peptide of BK strain hepatitis C
virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847[1998).

I Protein Sci. 7:837-847[1998).

- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMERANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral c precursor polyprotein, commonly with Asp or Glu in the P6
posttion, Cys or Thr in Pl and Ser or Ala in Pl.,
- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                              01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-52E-2003 (Rel. 24, Last annoration update)
65-67-67 (Rel. 42, Last annoration update)
66-67 (Greba) (Greba) (Greba) (Greba) (Fred. 67, Nonstructural protein R2 (Greba) 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDLINE-97015088; PubMed-8861916;
LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
MOOMBAW E.W., Adachl T., Hostomska Z.;
"The crystal structure of hepatitis C virus NS3 proteinase reveals a trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; Structure and organization of the hepatitis C virus genome isolated from human carriers.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-96235224; PubMed-8647104;
MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                              PRT; 3010 AA.
                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Cell 87:331-342(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus.
NCBI_TaxID=11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       {RNA}(N)
                                                                                                                                                                                                                                                                                                                                                              POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                            POLG_HCVBK
                                                                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                셤
                                                                                                  g
                                                                                                                                        ò
                                                               ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@iBb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN B (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/KELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
     BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                 LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REMOVED FROM CAPSID PROTEIN C BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PP01543; MCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS4a.
HCV_NS4b.
HCV_NS4b.
HCV_NS5a.
HCV_RGRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam; PF01538; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF010998; VIral_RFRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS2.
                                                                                                                                                                                                                                                           EMBL; M58335; AAA72945.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   115
191
383
383
729
1006
1615
2013
3010
369
1083
1107
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                              1A10; 25-MAR-98.
1JXP; 14-JAN-98.
1C2P; 15-NOV-09.
1CSJ; 08-NOV-99.
1GX5; 09-APR-02.
1GX6; 10-APR-02.
1GX6; 26-JUN-00.
80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002166;
InterPro; IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002518;
InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     116
192
384
730
1007
1007
11616
1863
2014
347
1083
1107
1107
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
                                                                                                                                                                                                                                                                                   A38465;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; InterPro; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                    PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
```

```
1064
                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                          Gaps
                      (POTENTIAL)
          (POTENTIAL).
                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                               Length 3010;
                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                   327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                               DB 1;
                                                                                                                                                                                                                                                                                                                          18;
                                                                                                                                                                                                                                                                                                              81.0%; Score 823.5; DB 1 76.5%; Pred. No. 1.6e-68;
                                                                        (GLCNAC.
(GLCNAC.
(GLCNAC.
                      GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
           (GLCNAC.
                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                          21; Mismatches
    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                              1185 CTRGVAKAVDEVPESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                          STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                          Conservative
                                                                             1120
1122
1133
1136
                                                                                                                                                                                                                                1157
1161
1163
1166
                                                                                                                                                                                                                                                                                              1688
                                                                                                                                                                                                                                                                                        1203 120
1680 168
3010 AA;
                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 156; Conserv
                                                                                                                                                                                                                                                      1168
1172
1175
1187
1189
                                                                                                                                                                                   1101
1104
1108
11120
11129
11139
                                                                                                                                                                   090
                                                                                                                                                                                                                                           1162
                                                                                                                     1039
1050
1059
1068
1068
1075
                                                                                                                                                      082
                                                                                                                                                                                                                                                                                              STRAND
                                       CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                          174
   CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                  STRAND
TURN
STRAND
STRAND
                                                                                                                                                                                              STRAND
STRAND
STRAND
STRAND
TURN
STRAND
HELLX
TURN
TURN
STRAND
                                                                                                                     HELIX
STRAND
STRAND
STRAND
                                                                                                                                                 STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                             STRAND
                                                                                                                STRAND
ð
                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                            ç
                                                                                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED cutstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                   01-AGG-1992 (Rel. 23, Created)
01-AGG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotations Capsian (Sporter of Proportion In Contains Envelope glycoprotein El (GPS2) (GPS2) (GPS2) (ROS2) (GPS2); Envelope glycoprotein El (GPS2) (GPS2) (ROS2) (ROS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [RNA](N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BY AN GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
STATLARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY $29.
                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Molecular cloning of the human hepatitis P;
proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D90208; BAA14233.1; -.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISCUSSION OF SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002519;
IPR002531;
IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR001410;
IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; A39253; GNWVCJ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11116;
POLG_HCVJA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
```

IPR002166;

œ

```
1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNOVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                   W POLYProtein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
W Core protein; Goat protein; Envelope protein; Helicase; ATP-binding;
W Transmembrane; Nonstructural protein; BY THE
CHAIN 116 191 MATRIX PROTEIN (POTENTIAL).
T CHAIN 192 383 MAJOR ENVELOPE PROTEIN E, POTENTIAL).
T CHAIN 194 NON-STRUCTURAL PROTEIN NS1 (POTENTIAL).
T CHAIN 1007 1615 PROTEASE/HELICASE NS3 (POTENTIAL).
T CHAIN 1661 1862 NON-STRUCTURAL PROTEIN NS4 (POTENTIAL).
T CHAIN 1663 2013 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
T CHAIN 1863 2013 RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRINLSGD-----TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC...) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                          RELAY SYSTEM (BY
RELAY SYSTEM (BY
RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81.0%; Score 823.5; DB 1; 75.5%; Pred. No. 1.6e-68; iive 23; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL.
CHARGE RELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
                                       Pfam; PPO1543; HCV_capsid; 1. Pfam; PPO1543; HCV_capsid; 1. Pfam; PPO1542; HCV_capsid; 1. Pfam; PPO1550; HCV_capsid; 1. Pfam; PPO1550; HCV_NS1; 1. Pfam; PPO1550; HCV_NS2; 1. Pfam; PPO1000; HCV_NS3; 1. Pfam; PPO1000; HCV_NS4; 1. Pfam; PPO1001; HCV_NS4; 1. Pfam; PPO0999; Viral_RdRP; 1. Probon; PP06602; HCV_NS1; 1. Probon; PD186062; HCV_NS1; 1. Probon; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 75.5
Matches 154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1125
                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED cutstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                         01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 23, Last sequence update)
01-A0G-1992 (Rel. 41, Last annotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RMa-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J8) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                    (RNA)(N).
-1. SUBUNIT. THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                     MEDLINE-9223022; PubMed-1314459; Okamoto K., Lizuka H., Tanaka T., Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T., Fukuda S., Tauda F., Mishiro S.; Fukuda S., Tauda F., Full-length sequence of a hepetitis C virus genome having poor homology to reported isolates; comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D10988; BAA01761.1; -.
                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001490;
InterPro; IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PIR; A40250; GNWVJB.
HSSP; P27958; IHEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                          NCBI_TaxID-11115;
                               POLG_HCVJ8
                                                                                                                                                                                                                                                                                                                                                                                                          genotypes.
RESULT 7
POLG_HCVJ8
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTOMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADV1PVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLMTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                           CELLULAR AMINOPEPTIONSE.
CAPSID PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Gaps
                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Bnvelope protein; Helicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INII_MET
                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    66.4%; Score 675; DB 1; 69.8%; Pred. No. 1.3e-54;
                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED GLCNAC
                                                                                                                                                                                                                                                                                                                                                        (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (
N-LINKED (
N-LINKED (
N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (
                                                                                                                                                                                                                                                                             POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             330177 MW;
Pfam: PF01539; HCV_env: 1.
Pfam: PP0150; HCV_NS1: 1.
Pfam: PF01538; HCV_NS2: 1.
Pfam: PF012907; HCV_NS3: 1.
Pfam: PF01006; HCV_NS4s: 1.
Pfam: PF01001; HCV_NS4s: 1.
Pfam: PF01506; HCV_NS5s: 1.
Pfam: PF01506; HCV_NS5s: 1.
ProDom: P1080662; HCV_NS1: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            69.88;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1091
2038
2359
2811
                                                                                                                                                                                                                               1619
1866
2017
3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1034
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139
                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
Matches 12
                                                                                                                                                                                    CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                         CHAIN
   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its meeb by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                              MEDILINE-9204440: PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
Machida A., Miyakawa Y., Mayuul M.;
Machida A., Miyakawa Y., Mayuul M.;
Machida A., Miyakawa Y., Mayuul M.;
I man darier: comparison with reported isolates for conserved
T from a human carrier: comparison with reported isolates for conserved
T and divergent regions. "I some a human carrier: comparison with reported isolates for conserved
T of den. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 73:8697-2704(1991).
J. Gen. Virol. 73:8697-2704(1991).
J. Gen. Virol. 74:8697-2704(1991).
J. Gen. Virol. 75:8697-2704(1991).
J. Gen. 
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 24) Last sequence update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope djycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivith)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48);
Hepatitis C virus (isolate HC-36) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [RNA][W].

-1- SUBGUIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -! SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_RdRP.
Helicase_C.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303. HSSP; P27958; 1HEI. MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PF01543; HCV_capsid;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_core; 1
HCV_env; 1.
HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS2; 1
HCV_NS3; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002518;
IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR007095;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002868;
                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                NCBL_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01001;
PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEROPS; U39.00]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   fam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
```

HCV_NS5a; 1

Pfam;

PRT; 3033 AA

STANDARD;

POLG_HCVJ6 P26660;

200

POLG_HCVJ6

RESULT 8

us-09-965-594-18.rsp

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                              78
                                                                                                                                                                                                                                                                                                                                                                                                                                         139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                             TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                            PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                        Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

INIT_MET 1 1 REMOVED FROM CAPSID PROTEIN C BY THE
                         SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                             CELULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                  CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                               66.3%; Score 674; DB 1; Length 3033; 68.7%; Pred. No. 1.6e-54; Live 27; Mismatches 29; Indels
                                                                                                                                                                                                                                                                                                                                F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC. . .)
(GLCNAC. . .)
                                                                                                                                                                                                                                                              (GLCNAC. .
                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                          (POTENTIAL).
                                                                                                                                                                                          GLCNAC
                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                   (GICNAC
                                                                                                                                            POTENTIAL.
                                                                                                                                                                                         N-LINKED
                                                                                                                                                                                DECH BOX
Pfam; PF00271; helicase_C; 1. Pfam; PF00998; Viral_RdRP; 1. ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                329165
                                                                                                                                                                                                                                                                                                                                                      Local Similarity 68.7
                                                                                                                                                                                                                                                                                                  649
1091
2038
2811
                                                                                                                                                                                                                                                                                                  649 64
1091 109
2038 20
2811 28
3033 AA;
                                                                                                                                                   ACT_SITE
ACT_SITE
                                                                                                                                                                 ACT_SITE
NP_BIND
SITE
                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
TRANSMEM
                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HHOA ARATH
                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                             ö
```

Arabidopsis thaliana (Mouse-ear cress). Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                     **RETAILS FROM NA.**

**RETAILS COLUMBAS.**

                                                                                                            SEQUENCE FROM N.A.
Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease, a homolog to the bacterial periplasmic protease HhoA.";
Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320. Schübert M., Peterson U., Funk C., Haas B., Schroeder W.P., Kieselbach T.;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Nature 402:769-777(1999).
                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 thaliana
```

opportunistic pathogen.";

```
use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   95 KTSPSVVYIEAIELPKTSSGDIL------ 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          140 ----SLLSPRPISYLK-----GSSGGPLLCPAGHAVGIFRAAVSTR--GVAKAV 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SGLGREIPSPNGKSISEAIQTDADINSGNSGGPLLDSYGHTIGVNTATFTRKGSGMSSGV 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182 --DPDNDLAVLKIETEGRELNPVVLGTSNDLRVGOSCFAI----GNPYGYENTLTIGVV 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2 KKKGSVVIVGRINL---SGDIAYAQQIRGEEGCQEISQIGRDKNQVEGEVQIVSTAIQI 57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58 FLATSINGVLW-----TVYH-------GAGTRTIASPKGPVTQM 89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                90 YINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-AICC 15692 / PAO1;
STRAIN-AICC 15692 / PAO1;
MEDLINE-20437337; PubMed-10984043;
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M. Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Santh K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saler M.H., Hancock R.E.W., Lory S., Olson M.V.; *Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                           Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide.
TRANSIT
CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68; Indels 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8.6%; Score 87; DB 1; Length 321; 21.3%; Pred. No. 0.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 16-OCT-2001 (Rel. 40, Created)
116-OCT-2001 (Rel. 40, Last sequence update)
116-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          209 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEASE HHOA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54; Conservative 31; Mismatches
                                                                                                                                                                                                                                                          MEROPS; S01.279; ...
InterPro; IPR001940; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00089; trypsin; 1.
PRINTS; PR00834; PROTEASFS2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                THYLAKOID
                                                                                                                                                             EMBL; AF114386; AAF24060.1; -.
EMBL; AL021710; CAA16717.1; ALT_SEQ.
EMBL; AL161548; CAB78839.1; ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        321 AA; 34691 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DF-IPVESLETIM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :| ||:::: |:
NFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID-287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAAD_PSEAE
Q9HX08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSIT
TRANSIT
CHAIN
DOMAIN
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1D DT TO DT 
   ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation-the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    138 RGSLLSPR--PIS-----YLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            92 NVDKDLVGWQAPQGSRSLTP-----CTCGSSDL-----YLVTRHADVIPVRRGDS 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81 QND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALKER---- 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   43 OVEGEVQ-IVSTATQIFLATSINGVL------WTVYHGAGIRTIASPKGPVTQMYT 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEGLARATH STANDARD; PRT; 437 AA.

022609; Q9LK85;
16-007-2001 (Rel. 40, Created)
16-007-2001 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Portease Do-like 1, chloroplast precursor (EC 3.4.21.-).
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
DEGPI OR DEGP OR AT3627925 OR K16N12.18.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29 QEEREVHFLISKAAQLVMATETDVALPAKPQAMQAFLTEYCGAAAGQI------RVFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-CV. Columbia;

MEDLINE-20363099; PubMed-10907853;

Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;

Kaneko T., Katoh T., Sato Rabidopsis thaliana chromosome 3. II.

"Structural analysis of Arabidopsis thaliana chromosome 3. II.

Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
Identification and characterization of Degp, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                   Nature 406:959-964(2000).
-I. SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE /
PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro: IPR003382; Flavoprotein.
Pfam: PF02441; Flavoprotein; 1.
Hypothetical protein: Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MM; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.4%; Score 85.5; DB 1; Length 209; 27.9%; Pred. No. 0.64; vative 16; Mismatches 61; Indels 5
                                                                                                                                                                                                                                                                                                                                                                                                            or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity 27.9% tes 51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PIR: H83144; H83144.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [3]
SEQUENCE OF 104-118.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   190 LET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LNT 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           and BAC clones
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
TO DETAIL THE SECOND SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
```

```
AAMP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 103 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
STRAIN-cv. Columbia;
Kieselbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SMISS-PROT data bank.
-!- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF.
                                                                                                                                                                                                                                                                                                                                                 Transit peptide; Chloroplast; Thylakoid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        148 SYL------KGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8.2%; Score 83.5; DB 1; Length 437; 26.2%; Pred. No. 2.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                          V -> 1 (IN REF. 2).
P -> S (IN REF. 2).
G -> R (IN REF. 2).
G -> D (IN REF. 2).
LL -> HF (IN REF. 2).
L -> V (IN REF. 2).
I -> V (IN REF. 2).
O -> E (IN REF. 2).
W; 1497Blab3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                          CHLOROPLAST (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55;
                                                                                                                                                                                                                                                                                                                                                                             PROTEASE DO-LIKE 1.
                                                                                                    -!- INDUCTION: By heat shock.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-!- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                       SERINE PROTEASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative 17; Mismatches
                                                                                                                                                                                                                                                     MEROPS; S01.279; ...
InterPro; IPR001478; PD2.
InterPro; IPR00140; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00089; rD2; 1.
PRINTS; PR00884; PROTEASES2C.
SMART; SM00228; PD2; 1.
                                                                                                                                                                                                                         EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            46213 MW;
                                                     CAN DEGRADE BETA-CASEIN.
                                                                                                                                                                                                                                                                                                                                                 Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                      321
421
171
201
280
23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local Similarity
Thes 45; Conserve
                                                                        O-PHENANTHROLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                            36
54
60
64
68
358
381
381
416
437 AA;
                                                                                                                                                                                                                                                                                                                                      PROSITE; PS50106;
                                                               ENZYME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                   TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                           RANSIT
  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

RESULT 12

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL cutstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           68 WIVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                123 TRHADVIPVRRR---GDSRGS-----LLSPRPISYLKGSSG--GPLLCPA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                           Homo sapiens (Human).
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----GHAVGIFR----AAVSTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8.2%; Score 83; DB 1; Length 452; 25.3%; Pred. No. 2.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative 13; Mismatches 47; Indels
                                                                                                                                                                                                                                                                                                                  Beckner M.E., Krutzsch H.C., Stracke M.L., Williams S.T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HEPARIN-BINDING (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MIM: 603488; -. GO: GO:0008201; F:heparin binding activity; TAS. InterPro; IPR001680; WD40.
                                      15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Anglo-associated migratory cell protein.
  452 AA
                                                                                                                                                                                                                                                        SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PROMIT PEOGROUP WD40; 8.
SMART; SM00320; WD40; 8.
PROSITE; PS00678; WD_REPEATS_1; 1.
PROSITE; PS50082; WD_REPEATS_2; 6.
PROSITE; PS5024; WD_REPEATS_REGION; 1.
Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-GLU.
WD 1.
WD 2.
WD 3.
WD 3.
WD 5.
WD 5.
WD 5.
                                                                                                                                                                                                                                                                            IISSUE-Brain;
MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M95627; AAA68889.1; -.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18
77
138
180
220
220
261
306
363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR; I39383; I39383.
Genew; HGNC:18; AAMP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
nes 42; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   452 AA;
                                                                                                                                                                                                               NCBI_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Baloinformatics and the EMBL outstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-92263794; PubMed-1316684;

Lacono-Connors L.C., Schmaljohn C.S.;

Lacono-Connors L.C., Schmaljohn C.S.;

Cloning and sequence analysis of the genes encoding the nonstructural proteins of Langat virus and comparative analysis with other flaviviruses.

1. CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.

-1. CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                      01-RBF.1993 (Rel. 25, Created)
01-FBF.1994 (Rel. 26, Last sequence update)
28-FBF.2003 (Rel. 41, Last annotation update)
28-FBF.2003 (Rel. 41, Last annotation update)
66-FBF.2003 (Rel. 41, Last annotation update)
Envelope protein PRM: Matrix protein (Envelope protein M): Major
envelope protein PRM: Matrix protein (Envelope protein M): Major
envelope protein PRM: Matrix Protein (Envelope protein M): Major
NS2A; Nonstructural protein NS2B; Helicase/protease (EC 3.4.21.98)
(NS3); Nonstructural protein NS2A; Nonstructural protein NS4B; RNA-
directed RNA polymerase (EC 2.7.7.48) (NS5)].
Langat virus (strain TP21)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 1-776 FROM N.A. SEQUENCE OF 1-776 FROM N.A. MEDILNE-92074260; PubMed-1720591; MANDLINE-92074260; PubMed-1720591; Mandl C.W., Iacono-Connors L., Wallner G., Holzmann H., Kunz C., Heinz F.X.; "Sequence of the genes encoding the structural proteins of the low-virulence tick-borne flaviviruses Langat TP21 and Yelantsev."; Virology 185:891-895(1991).
296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                           PRT; 3414 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Flavi_capsidc.
Flavi_glycoprotE.
Flavi_helicase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; M73835; AAA02740.1; ALT_TERM.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPRO00165; Flavi M.
IPR001157; Flavi NS2A.
IPR000752; Flavi NS2A.
IPR000487; Flavi NS2B.
IPR000404; Flavi NS4A.
IPR0001528; Flavi NS4B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002535; Flavi_propep
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FtsJ.
Hellcase_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; S35365; AAB22165.1;
PIR; A42545; A42545.
                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410;
InterPro; IPR001122;
InterPro; IPR000336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR001850;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1SVB
                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=31638;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S07.001;
                                                   RESULT 13
POLG_LANVT
ID POLG_LANVT
AC P29837;
                                                                                                                                                                                                                                                                                                                               Flavivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                               a
```

```
1496 GCSEGRSDSRPLDVKNGVYRIYTPGLLWGQRQIGVGYGAKGVLHTWHVTRGAALLVDGV 1555
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82 PKGPVTQMYINVDKDLV------GWOA-----PQGSRSLTPCTCGSSDLYLVT 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30 GCQETSOTGRDKNQVEGEVQIVSTA----TQTFLATSINGVLWTVYH---GAGTRTIAS 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLYPICTEUR (SUPPORTOTE): Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Envelope protein; Hydrolase; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Helicase; NNT_MET 1 REWOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.

CHAIN 113 205 CARSID PROTEIN C (POTENTIAL).

CHAIN 206 280 ENVELOPE PROTEIN M (POTENTIAL).

CHAIN 777 1128 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

CHAIN 1129 1358 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).

CHAIN 1139 1489 NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROTEASE/HELICASE (POTENTIAL).
NONSTRUCTURAL PROTEIN NS44 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLIMERASE (NS5)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----ENGREMGAI----PIDLAKGTSGSPIMNSQGEVVGLYGNGLKT 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
POTENTIAL.
BY SIMILARITY.
N-LINKED (GLCNAC. . . ) (POTENTIAL).
N-LINKED (GLCNAC. . . ) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              124 RHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVST 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 1; Length 3414;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC. , .) (POT. MW; 59CB7E95DD70D82E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL).
ATP (POTENTIAL).
DEAH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT; 3414 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7.8%; Score 79.5; D
22.7%; Pred. No. 69;
ive 21; Mismatches
                                                                                                                                                          Pram; Pr01004; Plavi_M; 1.
Pram; Pr01004; Plavi_M; 1.
Pram; Pr01005; Plavi_NS1; 1.
Pram; Pr01005; Plavi_NS2; 1.
Pram; Pr01340; Plavi_NS2B; 1.
Pram; Pr01349; Plavi_NS4; 1.
Pram; Pr01349; Plavi_NS4; 1.
Pram; Pr01349; Plavi_NS5; 1.
Pram; Pr01728; Plavi_Dropep; 1.
Pram; Pr01728; FrsJ; 1.
ProDom; Pr01728; PrsJ; 1.
ProDom; Pr001556; Flavi_glycoprote; 1.
ProDom; Pr00156; Plavi_glycoprote; 1.
ProDom; Pr001496; Flavi_NS1; 1.
Interpro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PP01003; Flavi_capsid; 1.
Pfam; PP02832; Flavi_glycop_C; 1.
Pfam; PP00869; Flavi_glycoprot; 1.
Pfam; PP00949; Flavi_halicase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    378017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SMART; SM00490; HELICC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           434 43
3414 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
es 39; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1129
1359
1490
2111
2260
2512
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1688
1779
103
262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_TBEVW
ID POLG_TBEVW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
TRANSMEM
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NP_BIND
SITE
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 14
     SOLITATION OF THE PROPERTY OF 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

æ

```
Flavi_glycoprotE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
DOMAIN
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
PROPEP
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HELIX
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its modified an or profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-Neudoerfl;
MEDLINE-96036491; PubMed-7483260;
Walliner G., Handl C.W., Kunz C., Heinz F.X.;
"The flavivirus 3' noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-Neudoerfl:
MEDLINE-88322870; PubMed-3413985;
Mandl C.W., Heinz F.X., Kunz C.;
"Sequence of the structural proteins of tick-borne encephalitis virus (western subtype) and comparative analysis with other flaviviruses.";
Virology 166:197-205(1988).
                   01-JAN-1990 (Rel. 13, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
66-06-06-07 (Rel. 41, Last annotation update)
protein (Rel. 41, Last annotation update)
protein (Core protein M); Major envelope protein E; Nonstructural proteins N31, NS2A, NS2B, NS4A and NS4B; Protease/helicase (EC 3.4.21.98) (NS3); RNA-directed RNA polymerase (EC 2.7.7.48)
(NS5)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 281-680.
MEDLINE-95272700; PubMed-7753193;
Rey F.A., Heinz F.X., Mandl C.W., Kunz C., Harrison S.C.;
"The envelope glycoprotein from tick-borne encephalitis virus at 2-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROYER BY WENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Neudoerfl;
MEDLINE-90051080; PubMed-2554575;
Mandl C.W., Heinz F.X., Stoeckl E., Kunz C.;
"Genome sequence of tick-borne encephalitis virus (Western subtype)
and comparative analysis of nonstructural proteins with other
flaviviruses.";
                                                                                                                                                                                                                                Tick-borne encephalitis virus (Western subtype) (TBEV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S07.UPW; -
Interpro; IPR001410; DEAD.
Interpro; IPR001122; Flavi_capsidc.
                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE OF 767-3414 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; U27495; AAA86870.1; -.
PIR; A31052; GNWVNE.
PDB; 1SVB; 10-JUN-96.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Virology 173:291-301(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 1-779 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    encephalitis virus.";
Virology 213:169-178(1995).
                                                                                                                                                                                                                                                                                                                    WCBI_TaxID=11088;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         resolution."
                                                                                                                                                                                                                                                                                           Flavivirus.
```

A PER PROPERTY AND PROPERTY AND

```
HYDROPHOBIC SIGNAL SEQUENCE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Prodom; PR001556; Flavi_10; ProprotE; 1.
Prodom; PD001496; Flavi_NS1; 1.
SWART; SW004496; Flavi_NS1; 1.
SWART; SW00490; HELICC; 1.
POlyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Envelope protein; Hydrolase; Helicase; ATP-binding; Transmembrane; Nonstructural protein; 3D-structure.
INII_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CHAIN 1 12 CAPSID PROTEIN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA-DIRECTED RNA POLYMERASE (NS5).
INVOLVED IN FUSION.
ATP (POTENTIAL).
DEAH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MAJOR ENVELOPE PROTEIN E.
NONSTRUCTURAL PROTEIN NS.1.
NONSTRUCTURAL PROTEIN NS.2.
NONSTRUCTURAL PROTEIN NS.2.
PROTEASE/RELICASE (NS.3).
NONSTRUCTURAL PROTEIN NS.4.
NONSTRUCTURAL PROTEIN NS.4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ENVELOPE GLYCOPROTEIN M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
POTENTIAL.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
       InterPro; InterPro; InterPro; InterPro; InterPro; InterPro; InterPro; IPR001850; Flavi_helicase.
InterPro; IPR000185; Flavi_helicase.
InterPro; IPR000185; Flavi_helicase.
InterPro; IPR000464; Flavi_hela.
InterPro; IPR001528; Flavi_hela.
InterPro; IPR001528; Flavi_hela.
InterPro; IPR001289; Flavi_hela.
InterPro; IPR001289; Flavi_hela.
InterPro; IPR001859; Flavi_hela.
InterPro; IPR001859; Havi_helase_C.
InterPro; IPR001095; Havi_helase_C.
InterPro; IPR001094; RNA_pol_bS_Ps.
InterPro; IPR001094; RNA_pol_bS_Ps.
InterPro; IPR001094; Flavi_dlycopro; IPP01001; Flavi_dlycopro; IPP01001; Flavi_dlycopro; IPP01001; Flavi_helase; I.Pfam; PF01005; Flavi_helase; I.Pfam; PF010728; Flavi_helase; I.Pfam; PF010728; Flavi_helase; I.Pfam; PF010728; Flavi_helase_C; I.Pfam; PF010778; Flavi_helase_C; I.Pfam; PF010778; Flavi_helase_C; I.Pfam; PF0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TURN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                             TISSUE-EXTREMBLYONIC tissue;

MCDOWell K.J., Adams M.H., Baker C.B.; KDDBJ databases.
Submitted (MAR-1995) to the EMBL/GenBark/DDBJ databases.
-!- FUNCTION: TRANSFERRINS ARE IRON BINDING TRANSPORT PROTEINS WHICH
CAN BIND TWO ATOMS OF FERRIC IRON IN ASSOCIATION WITH THE BINDING
OF AN ANION, USCALLY BICARBONATE. IT IS RESPONSIBLE FOR THE
TRANSPORT OF IRON FROM SITES OF ABSORPTION AND HENE DECRADATION
TO THOSE OF STORAGE AND UTILIZATION. SERMY TRANSFERRIN MAY ALSO
HAVE A FURTHER ROLE IN STIMULATING CELL PROLIFERATION.
                     01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Serotransferrin precursor (Transferrin) (Siderophilin) (Beta-1-metal binding globulin).
                                                                              Equus caballus (Horse).
Bukaryota, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
NCBI_TaxID=9796;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Transport; Iron transport; Glycoprotein; Metal-binding; Repeat;
                                                                                                                                                                                                                                                                                    -1- SUBCNIT: Monomer.
-1- SUBCRILIAR LOCATION: Secreted.
-1- TISSUE SPECIFICITY: Expressed in liver: secreted in plasma.
-1- DOMAIN: COMPOSED OF TWO HOMOLOGOUS DOMAINS.
-1- SIMILARITY: BELONGS TO THE TRANSFERRIN FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BY SIMILARITY.
SEROTRANSFERRIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
    706 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY.
SIMILARITY.
                                                                                                                             SEQUENCE FROM N.A.
MEDLINE-9377958; PubMed-8504171;
Carpenter M.A., Broad T.E.;
"The CDNA sequence of horse transferrin.";
Biochim. Biophys. Acta 1173:230-232(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LARITY
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Interpro; IPR001156; Transferrin.
Pfam; PF00402; TRANSFERIN; 2.
PRINTS; PR0042; TRANSFERIN.
SMART; SM00094; TR_FER; 2.
PROSITE; PS00205; TRANSFERIN_1; 2.
PROSITE; PS00205; TRANSFERIN_2; 2.
PROSITE; PS00206; TRANSFERIN_2; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMI
     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M69020; AAA30958.1; -. EMBL; U21127; AAA63684.1; -. PIR; S33761, S33761. HSSP; P02787; 1A8E.
    STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRFE_HORSE
P27425;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DISCLETO
DISCLETO
DISCLETO
DISCLETO
DISCLETO
DISCLETO
DISCLETO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIGNAL
    81 SPKGPVTQMYTNVDKDLV------GWQA-----PQGSRSLTPCTCGSSDLYLV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1555 AVAGP---YWADVREDVVCYGGAWSLEEKWKGETVQVHAFPPG-RAHEVHQCQPGELILD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34 TSQTGRDKNQ-----VEGEVQIVSTA-----TQTFLATSINGVLWTVYH---GAGTRTIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              T------GRKLGAIPIDLVKGTSGSPILNAQGVVVGLYGNGLKT 1648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   123 TRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVST 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    50;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 79; DB 1; Length 3414; Pred. No. 77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match 7.8%; Score 79; DB Best Local Similarity 23.7%; Pred. No. 77; Matches 41; Conservative 22; Mismatches
  1611
STRAND
TOTRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TURN
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
```

RESULT 15 TRFE_HORSE

g

ð

g

ð

셤

à

```
T DISULFID 498 692 BY SIMILARITY.
T DISULFID 508 522 BY SIMILARITY.
T DISULFID 590 604 BY SIMILARITY.
T DISULFID 590 604 BY SIMILARITY.
T DISULFID 642 647 BY SIMILARITY.
T METAL 79 IRON 1 (BY SIMILARITY).
T METAL 209 209 IRON 1 (BY SIMILARITY).
T METAL 270 270 IRON 1 (BY SIMILARITY).
T METAL 413 413 413 413 IRON 2 (BY SIMILARITY).
T METAL 449 449 470 IRON 2 (BY SIMILARITY).
T METAL 544 544 IRON 2 (BY SIMILARITY).
T METAL 612 612 IRON 2 (BY SIMILARITY).
T METAL 649 449 IRON 2 (BY SIMILARITY).
T BINDING 480 480 ANION (POTENTIAL).
T BINDING 480 480 ANION (POTENTIAL).
T CARBOHYD 515 515 N'LINKED (GICNAC. . .) (POTENTIAL).
T CARBOHYD 515 515 N'LINKED (GICNAC. . .) (POTENTIAL).
```

Query Watch
7.7%; Score 78.5; DB 1; Length 706;
Best Local Similarity 23.8%; Pred. No. 13;
Matches 48; Conservative 22; Mismatches 81; Indels 51; Gaps

178 VAKAVDFIPVESLE--TTMRSP 197 | | | | :::| | |:| | 557 VA---FVKHQTVEQNTDGRNP 574 carch completed: August 30, 2003, 19:13:48
cb time : 11.7567 secs

Page

hepatitis chepatitis c

091rr9 09dte2 099au2 Q91rt8 070818

091rs1 091rq8 091rt1 091rt1 091rt5 091rs3 091rt6 091rs8 091rt8 091rs7 091rs7 091rs7 091rs7 091rs7 091rs7 091rs7 091rs7 091rs7

```
SEQUENCE FROM N.A.
MEDILINE_203344; PubMed=10864644;
MEDILINE_203344; PubMed=10864644;
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-dependent bowlne viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Pestivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Lai V.C.H., Hong 2.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL, AF268278; AAF82566.1; -.
HSSP: P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 4040 AA
                                                                                                                                                                                                                                                                                                                                                                                                           ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR000280; CDvir_endptseP80.
InterPro; IPR000140; WC_NS3.
InterPro; IPR001410; WC_NS3.
InterPro; IPR001410; WC_NS3.
InterPro; IPR001505; Holloase_C.
InterPro; IPR001005; Myb_DNA_binding.
InterPro; IPR001005; Nyb_DNA_binding.
InterPro; IPR001005; Nyb_DNA_binding.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
Fam; PF02907; HCV_NS3: 1.
Ffam; PF002911; helicase_C; 1.
Ffam; PF002711; helicase_C; 1.
091R83
091R08
091R08
091RR0
091RR0
091R83
091R83
091R84
091R84
091R85
091R85
091R85
091R85
091R85
091R87
                                                                                                                                                                                                                                                                                                                              070818
Q91RR9
Q9DTE2
Q99AU2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        091FH8 PRELIMINARY;
091FH8;
01-0CT-2000 (TEMBLTE1.15,
01-0CT-2000 (TEMBLTE1.15,
01-MAR-2003 (TEMBLTE1.23,
  Genome polyprotein.
Mucosal disease virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11099;
MEROPS; S31.001;
                                                                                                                                                                                                                                                               838.5
838.5
838.5
838.5
837.5
837.5
837.5
841
840.5
840
839.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
 091FH8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HERER REPORT OF THE PRESENCE OF THE PROPERTY OF THE PRESENCE OF THE PROPERTY O
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091fh mucosal dis
03679 hepatitis c
091f56 hepatitis c
091fe5 hepatitis c
003463 hepatitis c
036608 hepatitis c
099wx5 hepatitis c
091rr8 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr5 hepatitis c
091rr7 hepatitis c
091rr7 hepatitis c
091rr7 hepatitis c
                                                                                                   (without alignments)
1352,314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pred. No. 1s the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                       August 30, 2003, 19:00:22; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Description
                                                                                                                                        US-09-965-594-18
1017
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP
                                                                                                                                                                                                                                                               830525
          GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd
                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                    830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUMMARIES
                                                                                                                                                                                                                                                                                                                           Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 036608
036608
099WX5
099RW9
091RR5
091RR5
091RR2
091RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q91FH8
Q36579
Q81756
Q91FES
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_vertebrate:*
sp_unclassified:*
sp_rvirus:*
sp_bacteriap:*
                                                                                                                                                                                             BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                     sp_fungi:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                            sp_archea:*
sp_bacterla:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_plant: *
sp_rodent: *
sp_virus: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DB
                                                                                                                                                                                                                                                                                                                                                                                SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2436
3011
3011
3011
3011
3015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                872.5
858.5
854.5
854.5
854.5
854.5
851.5
851.5
                                                                                                                                                       Perfect score:
Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     849
847
847
847
847
846
846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score
                                                                OM protein
                                                                                                                                                                                                                                    Searched:
                                                                                                                                                                                                                                                                                                                                                                                Database
                                                                                        on:
```

~

```
PF01543;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    081756
                      Pfan;
Pfan;
Pfan;
Pfan;
Pfan;
Pfan;
Pfan;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 3
081756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SO WEEK WAY WEEK WAS A STREET OF THE STREET 
                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                       Ĥ
                                                                                                                                                                                                                                                                                                                                                                                                              62 SINGVLWTVYHGAGTRIIASPRGPVTQMYINVDKDLVGWQAPQGSRSLIPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                           10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCRITSLIGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                         5 GSVVIVGRINLSGD---TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLAT 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               S.M.
PRINTS; PR00729; CDVENDOPTASE.
SMART; SM00487; DEXDC; 1.
SWART; SM00490; HELICO; 1.
PROSITE; PS00037; MTB_1; 1.
PROSITE; PS0507; RDRP_POSITIVE; 1.
PROSITE; PS05507; RDRP_POSITIVE; 1.
PROSITE; PS0551; RDRP_LT2_2; 1.
ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; BRNA-directed RNA, Polymerase; Transferase.
SEQUENCE 4040 AA; 453073 MW; ADE87791D05589DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Transmission of hepatitis C by intrahepatic inoculation with transcribed RNA.";
Science 277:570-574(1997).
Science 277:570-574(1997).
ILPOPMORIN: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPMORIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN MAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MINA (BY SIMILARITY).
                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-97373636; PubMed-9228008;
Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone
Rice C.M.;
                                                                                                                                                                                                                                       3.
                                                                                                                                                                                                   Length 4040;
                                                                                                                                                                                                                                       5; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last sequence update)
Last annotation update)
                                                                                                                                                                                                   Score 872.5; DB 12;
Pred. No. 1e-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT; 3011 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NSSA.
HCV_RdRP.
Helicase_C.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JAN-1998 (TrEMBLrel. 05, 01-JAN-1998 (TrEMBLrel. 05, 01-MAR-2003 (TrEMBLrel. 23, Genome polyprotein.
                                                                                                                                                                                                 Query Match
Best Local Similarity 89.2%;
Matches 174; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    182 VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       UDFIPVENLETTES 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AF009606; AAB66324.1;
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR001650; I
IPR007095; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002522;
InterPro; IPR002521;
InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002519;
IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            036579
036579;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DR DR KW
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
```

```
Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1065 ATGTFLATCINGVCMTVYHGAGTRTIASPRGPVIQMYINVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9; Gaps
                                                                                                                                                                                                                                                                                                                                      PROSITE; PSSOSO7; RDRP_POSITIVE; 1.
PROSITE; PSSOS21; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327182 MW; E2E0EE809C63C1B9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Choo Q.-L., Richman K., Han J.;
"The nucleotide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL; M32084; AAA456777.1; -.
HSSP: P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match 84.4%; Score 858.5; DB 12; Length Best Local Similarity 82.8%; Pred. No. 1.6e-74; Matches 169; Conservative 10; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TIEMBLIEL. 01, Created)
01-NOV-1996 (TIEMBLIEL. 01, Last sequence update)
101-NAR-2003 (TIEMBLIEL. 23, Last annotation update)
Genome POLyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 2436 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR000409; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001868; HCV_NS5A.
InterPro; IPR001866; HCV_RGAP.
InterPro; IPR001566; HCV_RGAP.
InterPro; IPR0010509; HA_POl_DS_PS.
InterPro; IPR001094; RNA_POl_DS_PS.
InterPro; IPR001094; RNA_POl_PSV1r.
Pfam; PP01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETIMRSP 197
                             R Pfam; PF01560; HCV_MS1; 1.

R Pfam; PF01560; HCV_MS1; 1.

R Pfam; PF01060; HCV_MS2; 1.

R Pfam; PF01000; HCV_MS4a; 1.

R Pfam; PF01001; HCV_MS4b; 1.

R Pfam; PF01506; HCV_MS5b; 1.

R Pfam; PF00271; Helicase_C; 1.

R Pfam; PF0099; Viral_RRRP; 1.

R PRODOM; PD186662; HCV_MS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
PF01542;
```

```
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Desai S.M., Devare S., Yamaguchi J.;
"Hepatitis C Virus.";
Submitted (ULL-200) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND WRNA (BY SIMILARITY).
EMBL; AE209078; AAG02099.1; -.
HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE; PS00190; CYPCCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
ATP-binding; Coat protein; protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ouery Match 84.0%; Score 854.5; DB 12; Length Best Local Similarity 82.8%; Pred. No. 3.9e-74; Matches 169; Conservative 9; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 3011 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAR-2001 (TrEMBLrel. 16, Created) 01-MAR-2001 (TrEMBLrel. 16, Last sequ 01-MAR-2003 (TrEMBLrel. 23, Last and Genome polyprotein.
                            InterPro; IPR002166; HCV_RdRP.
InterPro; IPR001659; Helicase_C.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF011543; HCV_capsid; 1.
Pfam; PF011543; HCV_care: 1.
Pfam; PF011569; HCV_env; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                           Pfam; PF01538; HCV_NS2; 1.
Pfam; PF012907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01506; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00971; helicase_C; 1.
Pfam; PF00998; viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEXDc;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-colonel;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 5
Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SO SE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ::| :::| : | :::| 55 RRGREILLGPADGNVSKGWRILAPITAYAQQTRGLLGCITSLTGRDKNQVEGEVQIVST 614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Gen. Virol. 82:1291-1297(2001).
-!- SUBUNIT: THE VIRIANO OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL: AF271632; AAF81759.1;
-HSSP: P271958; LALV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-21262212; Pubmed-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.;
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9; Gaps
                                                                                                                                                                                                                                                                             Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             84.0%; Score 854.5; DB 12; Length 2436; 82.8%; Pred. No. 2.9e-74; Live 9; Mismatches 17; Indels 9;
                                                                                                                                                                                                            PROSITE; PS50507; RDRP_VOSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
FYP-bInding; Coat protein; Envelope protein; Glycoprotein; Hydrolase; Nonstructural protein; Polyprotein; RNA directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                 2436 2436
2436 AA; 264734 MW; D7B9872900BF3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9IFE5;
01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HSSF; 72/353 LAAV.
INTERPRO; 1PR000345; CytC_heme_bind.
INTERPRO; 1PR001410; DEAD.
INTERPRO; 1PR001521; HCV_copsid.
INTERPRO; 1PR002521; HCV_core.
INTERPRO; 1PR002531; HCV_env.
INTERPRO; 1PR002531; HCV_NS1.
INTERPRO; 1PR002531; HCV_NS2.
INTERPRO; 1PR0004109; HCV_NS3.
INTERPRO; 1PR000145; HCV_NS4.
INTERPRO; 1PR001490; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     735 CTRGVAKAVDFIPVENLETTMRSP 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174 STRGVAKAVDFIPVESLETTMRSP 197
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicasc_C: 1.
Pfam; PF0098; Viral_RGRP; 1.
ProDom; P186062; HCV_NS1; 1.
SWART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sequence."
                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9IFE5
```

a õ

ô 셤 ò a 53

SON KKE VENE OF REAL PROPERTY OF REAL PR

```
5'-terminal sequence of the hepatitis C virus genome."; J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART: SM00487; DEXDC: 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002166; HCV_RGRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR001095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
Pfam; PF01543; HCV_Capsid; I.
                                                     STRAIN-HC-J1;
MEDLINE-92044440; PubMed-1658196;
                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-94174722; PubMed-7510436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5s; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; viral_RdRP; 1.
Probom; P1186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_env;
HCV_NS1;
HCV_NS2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002519; IPR002531; IPR002518; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR002522;
IPR002521;
                                                                                                                                                             [3]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PF01539;
                                                                                                                                                                                                                                                                                                    STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                     STRAIN-HC-J1;
                                                                                                                                                                                         STRAIN-HC-J1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pťam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
  54 AIQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                             protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepativirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-HC-J1;
MEDLINE-91013116; PubMed-2170712;
Okamoto H., Okada S., Suglyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           84.0%; Score 854.5; DB 12; Length
82.4%; Pred. No. 3.9e-74;
.1ve 11; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                          ATP-binding; Coat protein; Envelope protein; Glycoprotei Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transemembrane. SEQUENCE 3011 AA; 327107 MW; A6BECF5A3B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q03463 PRELIMINARY; PRT; 3011 AA.
003463;
01NOV-1996 (TIEMBLED. 01, Created)
01-NOV-1996 (TIEMBLED. 01, Last sequence update)
01-NOV-2003 (TIEMBLED. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Interpro; IPR002166; HCV_GRP.
Interpro; IPR001650; Helicase_C.
Interpro; IPR007095; NA_DOL_DS_PS.
Interpro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_copsid; 1
Pfam; PF01542; HCV_core; 1.
                                                                                                                                                                                                                                                                                                                        ProDom, PD186062; HCV_NS1; 1.
SMART: SMO0487; DEXDC: 1.
PROSITE; PS00190; CYTOCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITYE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
HCV_capsid
                                                                                                                                                                                                 Pram; PF01539; HCV_GCV; 1.
Pram; PF01539; HCV_GCV; 1.
Pram; PF01538; HCV_NS1; 1.
Pram; PF02907; HCV_NS3; 1.
Pram; PF012907; HCV_NS4; 1.
Pram; PF01001; HCV_NS4s; 1.
Pram; PF01506; HCV_NS5s; 1.
Promom; PF01506; HCV_NS1; 1.
                                                   HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                                                                                       HCV_NS5a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity 82.4%
Matches 168; Conservative
                                                                IPR004109;
IPR000745;
           IPR002521;
                        FPR002519;
                                                   IPR002518;
                                                                                            IPR001490;
                                                                                                         IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                   InterPro;
InterPro;
InterPro;
InterPro;
InterPro;
           InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
```

ç a ò g å g ठ g RESULT 6 003463

RANK RANK OCCOSE POTACL

virus isolated "Characterization and mapping of a B-cell immunogenic domain in hepatitis C virus E2 glycoprotein using a yeast peptide library."; virology 2010;246-255(1994).

-1- SUBDMIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL; D10749; BAA01582.1; -- HSSP; P27958; 1HEI. ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Okamoto H., Okada S., Sugiyama T., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.; Miyakawa Y., Mayumi M.; Mucleotide sequences of the genomic RNA of hepatitis C virus isolation a human carrier: comparison with reported isolates for conservand divergent regions.; J. Gen. Virol. 72:2697-2704(1991). MEDLINE-93117120; PubMed*1335573;
Okamcto H., Kanai N., Mishiro S.;
Frull-length nucleotide sequence of a Japanese hepatitis C virus
isolate (HC-11) with high homology to USA isolates.";
Nucleic Acids Res. 20:6410-6410{1992}. Mink M., Benichou S., Madaule P., Tiollais P., Prince A., Inchauspe G.; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane. SEQUENCE 3011 AA; 327112 MW; 97E9052C0250463B CRC64; Okamoto H.; Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases.

```
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR001490;
IPR002868;
IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP; P27958;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Repatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9PWX5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              O9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
g
                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                       ä
                                                                                                                                            54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                           114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-97385173; PubMed-9238047; Yanagal M.; Purcell R.H., Emerson S.U., Bukh J.; Yanagal M., Purcell R.H., Emerson S.U., Bukh J.; Transcripts from a single full-length cDNs clone of hepatitis C virus are infectious when directly transfected into the liver of a
                                                                                                                                                                                                                                           --- LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ⋖
                                                                     9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus strain H77.
Viruses: SRRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 12; Length 3011;
                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                     18;
                 Score 853.5; DB 1
Pred. No. 4.9e-74;
8; Mismatches 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410; DEAD.
InterPro; IPR001522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002531; HCV_MS1.
InterPro; IPR002531; HCV_MS2.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR004109; HCV_MS3.
InterPro; IPR001409; HCV_MS4.
InterPro; IPR001409; HCV_MS4.
InterPro; IPR001040; HCV_MS7.
InterPro; IPR001040; HCV_MS7.
InterPro; IPR001040; HCV_MS7.
INTERPO; IPR001040; HCV_MS7.
INTERPO; HCV_MS7.
INTERPORT INTERPROPERTY.
IPR001040; HCV_MS7.
INTERPORT INTERPROPERTY.
INTERPORT INTERPORTY.
INTERPORT INTERPROPERTY.
INTERPORT I
                                                                                                                                                                                                                                                                                                                                                                                                            174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            helicase_C; 1
                    83.9%;
82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS4a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS5a; 1
                                                                                                                     KKGSVVIVGRIN----
                                              cal Similarity 82.8
169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01506;
PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  chimpanzee.";
                    Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          036608;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
Pfam;
Pfam;
Pfam;
                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAPER REPORTED BY A STANKE STANKE SOLVE STANKE STANKE SOLVE 
                                                                                                                        õ
                                                                                                                                                                 셤
                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

```
1125 GGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLPRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                 3 KRGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BY A
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC: 1.
PROSITE: PS50507; RDRP_VORAL; 1.
PROSITE: PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 0875E6881CB5C198 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bukh J.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                             DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-99420396; PubMed-10489358;
Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.;
"Hepatitis C virus: an infectious molecular clone of a second igenotype (2a) and lack of viability of intertypic la and 2a chineras.";
                                                                                                                                                                                                                                                                          83.7%; Score 851.5; DB 12; Length
82.4%; Pred. No. 7.7e-74;
iive 10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_env.
HCV_NS1.
HCV_NS3.
HCV_NS4a.
HCV_NS4a.
HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AF177040; AAF01182.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4SSP; P27958; IHEI.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                       Local Similarity 82.49
```

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q91RR8
Q91RR8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100 091RRB 099 PRR 099 PRR 099 PRT 091 PRT 091
         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              පු
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1009 RRGQEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1068
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 ATQIFLATSINGVLWIVYHGAGTRIIASPKGPVIQMYINVDKDLVGWQAPQGSRSLIPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLRGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1069 ATQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Gaps
                                                                                                                                                                                                                                                                                                                                                                   PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
PROSITE; PS50507; RDRP_PS51TIVE; 1.
PROSITE; PS50507; RDRP_PS1TIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyproteein;
RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3015 AA; 328159 MM; B7D23BC1P190663A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-99420396; PubMed-10489358; Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.; Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.; Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic la and 2a chimeras.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 12; Length 3015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83.7%; Score 851.5; DB 12; Length 82.4%; Pred. No. 7.7e-74; Live 10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3015 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro: IPR001650; Helicase_C.
InterPro: IPR001219; Pyridoxal_deC.
InterPro: IPR001095; RNa_pol_DS_PS.
InterPro: IPR007094; RNa_pol_DS_PS.
InterPro: IPR007004; RNa_pol_DS_PS.
InterPro: IPR007004; RNa_pol_PSvir.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01509; HCV_NS2; 1.
Pfam; PF01001; HCV_NS2; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF001001; HCV_NS4; 1.
Pfam; PF00271; helicase_C: 1.
Pfam; PF00271; helicase_C: 1.
Pfam; PF00271; helicase_C: 1.
Pfam; PF00271; helicase_C: 1.
SVART; SM0487; DENDC: 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-MAK-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Virology 262:250-263(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ouery Match
Best Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepacivirus.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bukh J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        60Md60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 9
         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
1129 CGSSDLYLVTRHADVIPVRRRCDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Brvelope protein; Glycoprotein; Helicase; Hydrolase; Monstructural protein; Polyprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
NS4 protease (Fragment).
NS5 protease (Fragment).
NS6 protease (Fragment).
NS7 protease (Fragment).
NS8 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             83.7%; Score 851.5; DB 12; Length 3015;
1larity 82.4%; Pred. No. 7.7e-74;
Conservative 10; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hydrolase; Nonstructural process, conference RNA directed RNA polymerase; Transferase; Transmembrane enrinence 3015 AA; 328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AA
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CTRGVAKAVDFIPVENLGTTMRSP 1212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                    Interpro; IPR002518; HCV_NS2.
Interpro; IPR004103; HCV_NS3.
Interpro; IPR001409; HCV_NS4a.
Interpro; IPR001490; HCV_NS4b.
Interpro; IPR001266; HCV_RGRP.
Interpro; IPR001266; HCV_RGRP.
Interpro; IPR001569; Helicase_C.
Interpro; IPR001059; Pytdoxal_dc.
Interpro; IPR001095; RNA_pol_DS_PS.
Interpro; IPR007095; RNA_pol_DS_PS.
Interpro; IPR007095; RNA_pol_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF177039; AAF01181.1;
EMBL; AF177037; AAF01179.1;
                                                                                                                                                                                           Interpro; IPR001410; DEAD.
Interpro; IPR002522; HCV_C
Interpro; IPR002521; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEXDC;
                                                                                                                                                                                                                                                                                                                            Interpro; IPR002519; HC Interpro; IPR002531; HC Interpro; IPR002531; HC Interpro; IPR004109; HC Interpro; IPR001409; HC Interpro; IPR001409; HC Interpro; IPR001460; HC Interpro; IPR00166; HC Interpro; IPR00166; HC Interpro; IPR001650; HC INTERPRO; HC I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
Matches 168; Conserv
                                                                                                                                                      4SSP; P27958; 1HEI
```

```
79 IASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139 GSLLSPRPISTLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTWRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.; "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
                                                                                                                                                                                                                                                                     STRAIN-Pt.30;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (AFR-2011) to the EMBL/GenBank/DDBJ databases.
EMBL, AF36238; AR454563.1;
InterPro: JRR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                          NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepativirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369241; AAK54666.1; -
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 181
181 AA; 19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                181 181
181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.3%; Score 847; DB 12; 92.7%; Pred. No. 4.9e-75; ative 2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AA.
                                                                                                              Created)
                                                                                                           01-DEC-2001 (TrEMBLRel. 19, 01-DEC-2001 (TrEMBLRel. 19, 01-MAR-2003 (TrEMBLRel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                          [1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11103;
                                                                                                                                                                                                                         NCB1_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-Pt. 4V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                   Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                            RESULT 12
Q91RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 IASPKGPVIQMYINVDKDLVGWPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                   79 IASPKGPVTOMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                    64 IASPRGPVIQMYINVDKDLVGWPAPQGSRSLTPCTGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 TAYAQQTRGLLGCIITSLIGRDKNQVEGEYQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                       19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                        4 TAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYFGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                  139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                   124 GSLLSPRDISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                         "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus."; Clinical Strains of the Hepatitis C Virus."; Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

EMBL: AF369235; AAK54560.1; --
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TRAIN Pt.4:
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.
Submitted (APR-2001) to the EMBL/GenHank/DDBJ databases.
EMBL; AP369218; AR455543.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ;
0
                                                                                                                                                                                                                                                                       .;
0
                                                                                                                                                                                                                                   Score 849; DB 12; Length 181;
Pred. No. 3.1e-75;
1; Mismatches 11; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                          1 1
181 181
181 AA: 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 proclease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 849; DB 12;
Pred. No. 3.1e-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1: Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      83.5%;
93.3%;
                                                                                                                                                                                                                                        83.5%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    166; Conservative
                                                                                                                                                                                                                                                                       Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                          NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local S
                                                                                                                                                                                                                                          Match
                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q91RT5
Q91RT5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
õ
                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                     ò
```

Gaps

ø

```
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, RF802401, AAK54565.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3.
                             [1]
SEQUENCE FROM N.A.
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                               Protease.
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 IASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCICGSSDLYLVIRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                    79 IASPKGPVTOMYTNVDKDLVGWQAPQGSRSLIPCTCGSSOLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                         4 TATAQQTRGLLGCIXTSLIGRDKNQVEGEVQIVSTAXQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                              GSLLSPRPISYLKGSSGGPI.LCPAGHAVGIFRAAVSTRGVARAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 TAYAQQTRGEEGCQETSQTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-FLIC;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to Irls of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AFS69114; ARK5439.1;
InterPro; IPR004109; HCV_NS3.
Pfam: PF02907; HCV_NS3: I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q91RR3;
01-DEC-2001 (TFEMBLrel. 19, Created)
01-DEC-2001 (TFEMBLrel. 19, Last sequence update)
01-MAR-2003 (TFEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
NS3 protease (Fragment).
VIDER C vinus.
Hepatitis C vinus.
Hepatitis C vinus.
Hepativinus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                   .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  83.3%; Score 847; DB 12; Length 181; 92.7%; Pred. No. 4.9e-75;
                             Length 181;
                                                                                   11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 1
181 181
181 AA; 19114 MW; BEIDOBS42F014E86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q91RT9 PRELIMINARY; PRT; 181 AA.
Q91RT9;
G1-DEC-2001 (TIEMBLrel. 19, Created)
O1-DEC-2001 (TIEMBLrel. 19, Last sequence update)
O1-MAR-2003 (TIEMBLrel. 23, Last annotation update)
                    Score 847; DB 12;
Pred. No. 4.9e-75;
2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                          83.3%;
92.7%;
                    Query Match
Best Local Similarity 92.7%
Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 92.77
Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       091RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
Q91RT9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          091RR3
1D 099
DT 019
DT 010
DT 010
DT 010
DE NS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AND DESCRIPTION OF STATEMENT OF
                                                                                                                                              δ
                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                         ద
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                          d
```

```
ö
                                                                                                                                                                                                                                             64 IASPRGPVIQMYINVDRDLVGWPAPQGARSLTPCTGGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                          79 IASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                              4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                      139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                   ö
                                                             Length 181;
                                                         83.2%; Score 846; DB 12; Length 1 llarity 92.7%; Pred. No. 6.1e-75; Conservative 2; Mismatches 11; Indels
  181 181
181 AA; 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: August 30, 2003, 19:18:20 Job time : 38.5921 secs
                                                                            al Similarity
165; Conserv
NON_TER
SEQUENCE
                                                             Query Match
                                                                                    Local
```

```
AR145264 Sequence
AR145263 Sequence
AR145263 Sequence
AR145263 Sequence
AR145265 Sequence
AR145265 Sequence
AR145269 Sequence
AR145269 Sequence
AR145263 Sequence
AR145253 Sequence
AR145253 Sequence
AR145253 Sequence
AR145254 Sequence
AR145255 Sequence
AR145255 Sequence
AR145275 Sequence
AR145276 Sequence
AR145277 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AR110831 Sequence
BD069985 Functiona
AR118666 Sequence
106434 Sequence
109328 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AR118692 Sequence
106440 Sequence 54
109329 Sequence 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            109331 Sequence 15
M32084 Hepatitis (
AR118703 Sequence
                                                                                                                                                                                                                                                                                                                                                 AR179057 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR118696 Sequence
                                                                                                                                                                                                             Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALIGNMENTS
                                                                                                                                                                                                                                                                         SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AF009606
AR110831
BD069985
AR118686
IO6434
IO9328
                                                                                                                                                                                                                                                                                                                                                                              AR145268
AR145262
AR145263
AR145254
AR145266
AR145267
AR145261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ARI45270
ARI45274
ARI45272
ARI45255
ARI45257
ARI45257
ARI45251
ARI45271
ARI45271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145259
AR127810
BD081911
AR127809
BD081910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AR145258
AR145252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR118692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AR145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR14
em_vi:*
em_htg_hum:*
em_htg_inv:*
em_htg_other:*
em_htg_aus:*
                                                                                                                                                                                em_htgo_other:*
                                                                                                                                                   em_htgo_hum:*
                                                                                                                                                                   em_htgo_mus:*
                                                                         em_htg_pln:*
em_htg_rod:*
                                                                                                                        em_htg_vrt:*
                                                                                                          em_htg_mam:*
                                                                                                                                                                                                                                                                                                                      ВВ
                                                                                                                                                                                                                                                                                                    Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12980
12980
5360
5360
                                                                                                                                                                                                                                                                                                                                               RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Result
                                                                                                                                                                                                                                                                                                                      No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMMand line parameters:
-WODEL-frame+ p2n.model -DEV-x1p
-WODEL-frame+ p2n.model -DEV-x1p
-WODEL-frame+ p2n.model -DEV-x1p
-D-/cgn2_1/USPTO_spool/US0996594/runat_29082003_151919_28310/app_query.fasta_1.2872
-D9-/cgn2_1/USPTO_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-D9-Cabling-1-OFFMT-fastap -SUFFXT-rege -MIRMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-DOCALIGN-200 -THR_SCORE-pct -THR_MAXIN-0 -ALIGN-15 -MODE-LOCAL
-DOCALIGN-200 -THR_SCORE-pct -THR_MAXIN-0 -MAXIEN-2000000000
-USFR-050996594_GCGN1_1_14686_erunat_29082003_151919_28310_NCPU-6 -ICPU-3
-NO_MAAP -LARGEQUERY -NEG_SCORE-0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_IIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7
                                                                                           August 30, 2003, 19:18:33 ; Search time 2560.57 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                                                                                                                                            1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                         using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                               2888711 seqs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                     Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                           Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                          US-09-965-594-18
1017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              gb_ba:*
gb_htg:*
gb_in:*
gb_ov:*
gb_ov:*
gb_pat:*
gb_pl:*
gb_pl:*
gb_pr:*
gb_sts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_hum:*
em_in:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    em_ph:*
em_pl:*
em_ro:*
em_sts:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         em_pat:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                gb_un:*
gb_vi:*
em_ba:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_mu:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          em_or:*
                                                                         nucleic search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GenEmbl:*
                                                                                                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Title:
Perfect score:
                                                                                                                                                                                                                          Scoring table:
                                                                         OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                                                                 Searched:
                                                                                                      Run on:
```

us-09-965-594-18.rge

ACCESSION AR145264 VERSION AR145264.1 GI:1510 KEYWORDS OURCOWN. ORGANISM UNKNOWN.	onclassi onclassi Malcolm, Single-c protease Patent:	All a	Pred. No.: 1.1e-65 Score: 881.50 Percent Similarity: 93.378 Best Local Similarity: 85.20% Query Match: 6 DB: 6 US-09-965-594-18 (1-197) x AR14	Oy 5 GlySerValValleVa		н		Db 184 AACCAGGTCGAGGGAGAGA	Db 244 TGCGTCAACGCGTGTG OV 82 Professive From 171	6	0y 102 AlabroGlnGlySerAr ::: Db 364 GCGCCCCGGGGGCGC	Qy 122 ValThrargHisAlaAs	Db 424 GTCACGAGACATGCTG Qy 142 LeuSerProArgProI)	484	544 TCGGGGCACGCTGTGG	Oy 182 ValAspPheIleProVa	SULT 3 145268 CUS AR CUS AR FINITION SE CESSION AR
AR179057 AB179057 12734 bp DNA linear PAT 20-APR-2002 LOCUS DEFINITION Sequence 1 from patent US 6326137. ACCESSION AR179057 G1:20220612	SOURCE Unknown. ORGANISM Unknown. ORGANISM Unknown. Unclassified. NATHORS Hong, Z., Lai, V.C.H. and Lau, J.Y.N. TITLE Hepatitis C virus procease-dependent chimeric pestivirus JOURNAL Patent: US 6326137-A 1 Q4-DEC-2001; FEATURES	SOURCE 112/34 Aorganism="unknown" BASE COUNT 4032 a 2604 c 3295 g 2803 t ORIGIN	Alignment Scores: 1.11e-65 Length: 12734 Score: 892.50 Matches: 177 Percent Similarity: 92.87% Conservative: 4 Best Local Similarity: 90.77% Mismatches: 11 Ouery Match: 6 Gaps: 1	US-09-965-594-18 (1-197) x AR179057 (1-12734)	Oy S GlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAlaTyr 21	22 AladholnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys	473 GCCCAGCAGAGGGCCTCCTAGGGTGTAAGATCACCAGTCTGACTGGCCGGGAAA	UY 4. ANDANDANDELDAGIQUUARIQUILLEVALGETITATATITATGITITEPRELEMATIT 61 [HIH	Oy 62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81 	82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuvalGlyTrGln	Db 653 CCCANGGGTCTGTCATCCACATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCC 712 Qy 102 AlaProGlnGlySerArgScrLeuThrProCysThrCysGlySerScrAspLeuTyrLeu 121	Db 713 GCTCCTCAAGGTTCCCGCTCATTGACACCTGCGGCTCCTCGGGCTTTACCTG 772	Oy 122 ValThrargHisAlaAspValIleProValArgArgGlyAspSerArgGlySerLeu 141 	Oy 142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161	162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 18	Db 893 GCGGGACACGCCGTGGGCCCTATTCAGGCCCGCGGTGTGCACCCGTGGAGTGGCCAAGGCG 952 Ov 182 ValasphelleProValGluSerLeuGluThrThrMetArgSer 196	

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                        243
                                                                                                                                                                                                         31yGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                         21
                                                                                                                                                                         AlGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                               98)
mil.S.Shane., Weber, P.C. and Yao, N.
mbinant complexes of hepatitis C virus NS3
cofactor peptide
8-A 105 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                             1998
167
16
10
3
                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1998 bp DNA
Sequence 109 from patent US 6211338.
AR145268
AR145268.1 GI:15107135
                                                                                     ų
                                                                                     423
                                                                                                                                                           145264 (1-1998)
                                                                              m-"unknown"
c 569 g
                                                                Qualifiers
07131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
ACCESSION
VERSION
```

```
Query Match:
                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
ORGANISM
                                                                         BASE COUNT
                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                    AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
                                           JOURNAL
              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR145263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 5
                                                   FEATURES
                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
                                                                                                                                                                       ò
                                                                                                                                                                                     q
                                                                                                                                                                                                     à
                                                                                                                                                                                                                 đ
                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                 윱
                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-AUG-2001
                                                                                                                                                                                                                                                                                                                                      183
                                                                                                                                                                                                                                                              243
                                                                                                                                                                                                                                                                                    122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                          LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                              162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                             603
                                                                                                                                                                                                                   41
                                                                                                                                                                                                                                                61
                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                                                                      SerIleAsnGlyValLeuTrpThrValTyrH1sGlyAlaGlyThrArgThrIleAlaSer
                                   P.C. and Yao, N.
hepatitis C virus NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                   1998
166
17
10
3
                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                           1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.
Single-chain recombinant complexes of h
protease and NS4A cofactor peptide
Patent: US 6211338-A 109 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 103 from patent US 6211338.
                                                                                      423
                                                                                                                                                         Gaps:
                                                                                                                                                                       US-09-965-594-18 (1-197) x AR145268 (1-1998)
                                                                Location/Qualifiers
                                                                               /organism="unknown"
595 c 569 g
                                                                                        6
                                                                                      569
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145262.1 GI:15107129
                                                                                                                   1.99e-65
878.50
93.37%
84.69%
86.38%
                                                                        .1998
                     Unclassified
                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                       40
      Unknown
               Unknown
                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                            182
                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                             62
                                                                                                                                                                                                                                                                                                         82
                                                                                                                                                                                                                                                                                                                       304
                                                                                                                                                                                                                                                                                                                                                    364
                                                                                                                                                                                                                                                                                                                                                                                                               184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         504
                                                                                                                                        Best Local Si
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                         source
      SOURCE
ORGANISM
                                                                                      BASE COUNT
                                                                                                                     Pred. No.:
                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 4
AR145262
LOCUS
                                                         JOURNAL
                                                                FEATURES
                                           TITLE
                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                쉱
                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                요
                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                     ö
                                                                                                                                                                                                  g
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                Š
                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                           4
                                                                                                                                                                                                                                           GlyserValVallleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                                        AlaGinGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
      Unclassified.

1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 103 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                         1998
167
15
11
3
                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR145263 1998 bp Di
Seguence 104 from patent US 6211338.
AR145263
                                                                                                  424
                                                                                                                                                                                                             (1-1998)
                                                                                      /organism-"unknown"
596 c 568 q
                                                                                                                                                                                                              US-09-965-594-18 (1-197) x AR145262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GI:15107130
                                                                                                                                        2.42e-65
877.50
92.86%
85.20%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145263.1
                                                                                                                                                                       Best Local Similarity:
                                                                                                  410 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown.
Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown
                                                                                                                                                              Percent Similarity:
                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                 2
                                                                                                                                                                                                                                                                                                                                                         62
                                                                                                                                                                                                                                                                                                                                                                                                                  304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               424
                                                                                                                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       604
```

Unclassified

```
Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unknown
                                                                                                       Percent Similarity:
                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                           142
                                                                                                                        Ouery Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOURCE
ORGANISM
                                    source
                                                                                        ..
9
                                                     BASE COUNT
                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR145266
                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                               ORIGIN
                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                          合
                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               용
                                                                                                                                                                                 g
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                   õ
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                       64 GGTTCTGTTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGCCTAC 123
                                                                                                                                                                                                                  LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                      244 TGCGTCAACGCCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303
                                                                                                                                                                                                                                                                                    SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                  GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                    NS3
1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinat complexes of hepatitis C virus
protease and NS4A cofactor peptide
Patent: US 6211338-A 104 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases 1 to 651)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR145254 651 bp |
Sequence 95 from patent US 6211338.
AR145254
                                                                     424
                                                                                                                                                 Gaps:
                                                                                                                                                                 US-09-965-594-18 (1-197) x AR145263 (1-1998)
                                                            /organism="unknown"
596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR145254.1 GI:15107121
                                                                                                       2.42e-65
877.50
92.86%
85.20%
86.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unclassified
                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                                                          424
                                                                                                                                                                                                                                                                                                                                                                                                                           142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             604
                                                                                                                                                                                                                                                                                                                                                                                                                                           184
                                                                                                                               Best Local S
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145254
LOCUS
DEFINITION
ACCESSION
                                                      source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                   JOURNAL
          AUTHORS
                                            FEATURES
                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                  ô
                                                                                                                                                                                                  엄
                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                   ò
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      243
                                                                                                                                                                                                                                                                                                                                                                                                                                                      101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       603
                                                                                                                                                                                                                                                              303
                                                                                                                                                                                                                                                                                                                         124 TCCCAACAGACGCGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                                                                                                                                                                                  41
                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                    81
                                                                                                                                                                                                                                                 7
                                                                                                                                                                                                                                                                                                                                                                                                               62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                   42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (Dasses 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
  hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          651
166
10
3
                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Single-chain recombinant complexes of protease and NS4A cofactor peptide Patent: US 6211338-A 95 03-APR-2001; Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6211338
                                                                                                                                                                                 Indels:
Gaps:
                                                                               144
                                                                                                                                                                                                                     (1-651)
                                                             /organism-"unknown"
187 c 200 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR145266 1998
Sequence 107 from patent US
AR145266
AR145266.1 GI:15107133
                                                                                                                                                                                                                       US-09-965-594-18 (1-197) x AR145254
                                                                                                                             1.18e-65
874.50
93.33%
85.13%
85.99%
                                                   1. .651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unclassified
                                                                                                                                                                    Best Local Similarity:
                                                                             120 a
```

S

source 1, 1998	BASE COUNT 410 a 596 ORIGIN	ent Scores: No.:	Score: 874.5 Percent Similarity: 92.86 Best Local Similarity: 84.69 Query Match: 85.99 DB: 6	US-09-965-594-18 (1-197) x AR	Qy 5 GlySerValvalile	49 (22 Alacincininiary 23 Alacincininiary 24 TCCCAACAGGGGG	42	184	UY	83	::: Db	Qy 102 AlaProGlbGlySer	364		Db 424 GTCACGAGACATGCT Qy 142 LeuSerProArgPro	484	Qy 162 AlaGlyHisAlaVal	Db 544 regegeAcergre	Oy 182 ValAspPheIlePro	604	AR145261 LOCUS AR145261 LOCUS DEFINITION Sequence 102 from ACCESSION AR145261 VERSION AR145261 GI:15 KEYWORDS SOURCE UNKNOWN.		ą.
JOURNAL Patent: US 6211338-A 107 03-APR-2001;	source	N N N N N N N N N N N N N N N N N N N	Alignment Scores: 4.35e-65 Length: 1998 Sred. No.: 874.50 Matches: 166 Percent Similarity: 92.86% Conservative: 16 Best Local Similarity: 84.69% Mismatches: 11	Indels: Gaps:	US-09-965-594-18 (1-197) x AR145266 (1-1998)	5 GlySerValValleValGlyArglleAsnLeuSerGlyAspThrAlaTyr	DD 04 GGITCIGITGITATIGITGGGGGGGTTATITTTTGGGGGGGGGG	:::!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!	42 AsnGlnvalGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr	DD 184 AACCAGGIGGAGGAGAGGIICAGGIGGIITCACGGCAACACAACCTICCIGGCGACC 243 Ov 62 SertlakanglaVallantnThTVallacHisclutlaclutlachthtlablaSer 81	244 IGCGTCAACGCGTGTGTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTTAGCCGGC	Oy 82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101	Db 304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGGCTGGCAG 363	102 AlabroGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu	364 GCGCCCCCCGGGGCGCGTTCCTTGACATGCACCTGTGGGCAGCTCAGACCTTTACTTG	Oy 122 ValThrargHisAlaaspValleProValArgargargOlyassSerragGlyssereu 141	Qy 142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuJeuCysPro 161	Db 484 CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGCTCTGCTGGTCGACTGCTCTGCCCT 543	162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla	Db 544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGTTGCGAAGGCG 603	Oy 182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197 	RESULT 8 AR145267 LOCUS LOCUS DEFINITION Sequence 108 from patent US 6211338. ACCESSION AR145267 ACCESSION AR145267. GI:15107134 KEYWORDS	_	AUTHORS Malcolm, B.A., Taremi,S.Shane., Weber,P.C. and Yao,N. TITLE Single-chain recombinant complexes of hepatitis C virus NS3 protease and NS4A cofactor peptide JOURNAL Patent: US 6211338-A 108 03 APR-2001; FEATURES Location/Qualifiers

```
\text{alg1} \text{alg2} \text{
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         aremi, S.Shane., Weber, P.C. and Yao, N. combinant complexes of hepatitis C virus NS3 4A cofactor peptide 338-A 102 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     OVAlGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                  1998
166
11
11
                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1998 bp DNA
Om patent US 6211338.
                                                                        424
                                                                                                                                                                                                                                                                                                                                                                                                                                                 R145267 (1-1998)
98
ism-"unknown"
6 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                on/Qualifiers
98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sm- unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5107128
                                                                                                                                                                                                          .e-65
.50
.50
.99
```

```
VRL 12-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pestivirus type 1
Pestivirus type 1
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Pestivirus.
                                                                                                                                                                                                                                                                                                                                                           561
                                                                                                                                                                                                                                                                                              381
                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                          441
                                                                                                                                                                                                                                                                                                                                                                                                       161
                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                                                                                                                162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                   201
                                                                                                                                                                                                            261
                                                                                                                                                                                                                                                    321
                                                                                                                                                                                       61
                                                                                                                                                                                                                                 81
                                                                                                                                              22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                 142 TCCCAACAGACGCGGGCCTACTTGGTTGCATCATCACTAGCCTTACAGGCCGGGACAAG
                                                                                                                                                                                       42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                262 TGCGTCAACGGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                     102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Lal.V.C., Zhong.W., Skelton,A., Ingravallo,P., Vassilev,V., Donis,R.O., Hong.Z. and Lau.J.Y.
Donis,R.O., Hong.Z. and Lau.J.Y.
Donisser.O., Hong.Z. and Lau.J.Y.
Droftenser.O., Hong.Z. and Lau.J.Y.
Droftenser.O., Hong.Z. and Lau.J.Y.
J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
      2016
167
14
12
3
       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AF268278 12734 bp
pestivius type 1, complete genome.
AF268278.1 GI:9049956
                                                             Gaps:
                                                                                 US-09-965-594-18 (1-197) x AR145269 (1-2016)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2 (bases 1 to 12734)
Lai, V.C.H. and Hong, Z.
Direct Submission
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (bases 1 to 12734)
      5.35e-65
873.50
92.35%
85.20%
85.89%
                          Percent Similarity:
Best Local Similarity:
Query Match:
Alignment Scores:
                                                                                                                                                                                                                                  62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              622
                                                                                                                                                                                                                                                                                                                                        382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ద
                                                                                                                                                              g
                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                      δ
                                                                                                                       a
                                                                                                                                             ò
                                                                                                                                                                                      \delta
                                                                                                                                                                                                      a
```

JOURNAL

```
IIRYGLWGTHTALYKSIAARLGHETAFATLVIKWIAFGGESVSDHVKQAAVDLVYYYV
WNKPSFPQISETQQEGRRFYASLFISALATYTYKTWNYHNLSRVVEBALAYLFYATSA
WNKPSFPRDSSTQQEGRRFYASLFISALATYTYKTWNYHNLSRVVEBALAYLFYATSA
LKWFTPTRLESVVILSTTIYTYLSIRKGKSDGLLGTGISAAMILSRQNPSVGLSVM
LCVGAIAAHNAIESESECRRTLLMYRFYKNFLOQAATDELVKBNPERIIMALFERVQTI
GNPLRLIYHLYGVYYKGWEAKELSERTAGRNLFTLIMFEAFELLGMDSQGRIRNLSGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YILDLIYGLHKQINRGLKKWVLGWAPAPFSCDWTPSDERIRLPTDNYLRVETRCPCGY
EMRAFKNVGGKLTKVEESGPFLCRNRPGRGPVNYRVTKYYDDNLREIKPVAKLEGQVE
HYYKGVTAXIDYSKGKMLLATDKWEVEHGVITRLAKRYTGVGFNGRYLGDEPNHRALV
(16-MAY-2000) Antiviral Therapy, Schering-Plough Research, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA Location/Qualifiers
                                                                                                                                                                                                                                                                                                    /translation-*MELNTNEGSGSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSL
TYROKNQVBCEVQIVSTATOTFLATCINGVCMTVIHGAGTRITASPKGPVIONTINVD
ODLYGWPAPPGGSRSIIPCTGSSDLYLVTRHANTVPWRRGDSRGSLLSPRPISYLKG
SSGGPLLCPAGHAVGLRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
SYSDTKEEGATKKKTQKPDRLERGKMKIVPKESEKDSKTRPPDATIVVEGVKYQVRKK
                                                                                                                                                                                                                                                                                                                                                                                                                        GKTKSKNTQDGLYHNKNKPQESRKLEKALLAMAIIAIVLFOYTMGENITQWNLODNG
TRGIORAMAGNGVNRSLHGINPEKICTGVPSHLATDIELKTIGMUNDAEKTNYTGCR
LORHEMNKHGMCNWTNIEPGINPEKICTGOPSHLEGDPRECAYTCRYDRASDLNVYTQA
RDSPTPLTGCKKGGKWTNIEPGAILMNRTQONLTERQPRECAYTCRYDRASDLNVYTQA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VLEKVEPDASDNQNSVKIGLDEGNY PGPGIQTHTLTEEIHNRDARPFIM1LGSRNSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nraktarninlytgndpreirdlmaagrmlyvalrdydpelsemydfkgtfldreale
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ERDCATITKNTVOFLKMKKGCAFTYDLTISNLTRLIELVHRNNLEEKEIPTATVTTWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AYTFVNEDVGTIKPVLGERVIPDPVVDINLQPEVQVDTSEVGITIIGRETLMTTGVTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ALSLGOPRPROVIKEAVRNLIEQKKDVEIPNWFASDDPVFLEVALKNDRYYLVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               vkdoakalgatdotriikevgsrtyamklsswfloasnromsltplfeelllrcppat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        KSNKGHMASAYQLAQGNWEPLGCGVHLGTIPARRVKIHPYEAYLKLKDFIEEEEKKPR
                                                                                   /organism-"Pestivirus type 1"
/mol_type-"genomic RNA"
/db_xref-"taxon:11099"
                                                                                                                                                                                                                                    /product-"polyprotein"
/protein_id-"AAF82566.1"
/db_xref-"GI:9049957"
                                                                                                                                                                                                              /codon_start=1
                                                                                                                                                                                   .12508
                                                                                                                                                                    386
                        Institute,
Submitted
                                                                              source
                                                                                                                                                                  5'UTR
CDS
```

VKDTVIREHNKWILKKIRFQGNLNTKKMLNPGKLSEQLDREGRKRNIYNHQIGTIMSS

```
DITORMITEVEVITAGENT INTEGENCE AND SUBJECT CONTRIBUTION OF THE STRING STR
                                        PKNEKRDYSDDWQAGDLVVEKRPRVIQYPEAKTRLAITKVMYNWVKQQPVVIPGYEGK
TPLFNIFDKVRKEWDSFNEPVAVSFDTKAWDTQVTSKDLQLIGEIQKYYYKKEWHKFI
                        /GEVTWEQLEAGINRKGAAGFLEKKNIGEVLDSEKHLVEQLVRDLKAGRKIKYYETAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            08-AUG-2001
agirleklpivraqtdtktfheairdkidksenronpelhnklleifhtiaqptlkht
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGTAGTGTTGTTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTGC 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlyServalvalileValGlyArgileAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             651 bp
from patent US 6211338.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-18 (1-197) x AF268278 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GI:15107125
                                                                                                                                                                                                                                                                                                                  5.58e-64
872.50
91.79%
89.23%
85.79%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66
                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR145258.1
                                                                                                                                                                                                                                       ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR145258
Sequence 9
AR145258
                                                                                                                                                                                                                                     4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown
                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ŋ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                              3'UTR
BASE COUNT
                                                                                                                                                                                                                                                                                                                         ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
AR145258
                                                                                                                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       윱
```

```
Alignment Scores:
                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                        184
                                                                                                                                                                                                                                                                                      82
                                                                                                                                                                                                                                                                                                                    102
                                                                                                                                                                                                                                                                                                                                    364
                                                                                                                                                                                                                                                                                                                                                    122
                                                                                                                                                                                                                                                                                                                                                                  424
                                                                                                                                                                                                                                                                                                                                                                                   142
                                                                                                                                                                                                                                                                                                                                                                                                                 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
DEFINITION
                                                      BASE COUNT
ORIGIN
                                                                                      ..
No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
TITLE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
AR145253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                       JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KEYWORDS
SOURCE
                                FEATURES
                                                                                              Score:
                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                      a
                                                                                                                                                                                         ò
                                                                                                                                                                                                       엄
                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                            ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-AUG-2001
                                                                                                                                                                                                                                                                                                    243
                                                                                                                                                                                                                                                                                                                                                                   ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                        64 GGTTCTGTTGTTATTGTTGGTAGAATTATTTATCTGGTAGTGGTAGTATCACGCCTAC 123
                                                                                                                                                                                                                 41
                                                                                                                                                                          GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                       SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                        AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
      1 (bases 1 to 651)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus 1
protease and NS4A cofactor peptide
Patent: US 6211338-A 99 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              648
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        651
165
17
10
3
                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              51 bp 1
6211338.
                                                                      144
                                                                                                                                                           (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         651 <sup>3</sup>
US 6
                                                Location/Qualifiers
                                                             /organism="unknown"
187 c 200 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145252
Sequence 93 from patent
AR145252
                                                                                                                                                            US-09-965-594-18 (1-197) x AR145258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AR145252.1 GI:15107119
                                                                                                     2.12e-65
871.50
93.33%
84.62%
85.69%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unclassified.
1 (bases 1 to 651)
                                                     .651
                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                      120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown
                                                                                               Alignment Scores:
                                                                                                                                                                                                         22
                                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                       184
                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                                                                   102
                                                                                                                                                                                                                                                                                                                                                   364
                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                  434
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              604
                                                                                                                                  Query Match:
DB:
                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
               AUTHORS
                                                                      BASE COUNT
ORIGIN
                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                        JOURNAL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AR145252
                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                          õ
                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                  ç
                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                          õ
                                                                                                                                                                                         a
                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             161
                                                                                                                                                                                                                                                                                                      243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    423
                                                                                                                                                                                                                                                                                                                                                                                  AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                  41
                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         484 CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGCTCTTCGGGTGGTCCACTGCTCTGCCCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                                                                                                                                                                                                                                                                                                                                                                                                             ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                 22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                          AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 (bases 1 to 651)
Malcolm, B.A., Tarem1, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N. Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide Patent: US 6211338-A 93 03-APR-2001; Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               651
166
15
11
                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                             ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145253 651 bp
Sequence 94 from patent US 6211338.
AR145253. GI:15107120
                                                                                                           145
                                                                                                                                                                                                                                                                           US-09-965-594-18 (1-197) x AR145252 (1-651)
                                                                                          /organism="unknown"
188 c 199 g
                                                                                                                                                                    2.58e-65
870.50
92.82%
85.13%
85.59%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unknown.
Unclassified.
                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unknown.
```

σ

```
Percent Similarity:
                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                             122
                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      604
                                                                                                                                                                                                                 42
                                                                                                                                                                                                                                                   62
                                                                                                                                                                                                                                                                                                                         102
                                                                                                                                                                                                                                                                                                                                           364
                                                                                                                                                                                                                                                                                                                                                                              424
                                                                                                                                                                                                                                                                                                                                                                                                                 484
                                                                                                Query Match:
           source
                                                                :.
2
                           COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search cor
Job time
 FEATURES
                                     ORIGIN
                             BASE
                                                                Pred
                                                                                                                                                                                           g
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                      ď
                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                 ď
                                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                            ð
                                                                                                                                                            g
                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                  543
                                                                                                                                                                      243
                                                                                                                                                                                                                                                                                82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                         363
                                                                                                                                                                                                                                                                                                                                          AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                    ValThrargHisAlaaspValIleProValargArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                GTCACGAGACATGCTGACGTCATTCCGGTGCGCGGCGGGGGGGACAGTAGGGGGGAGCCTG 483
                                                                                                                                                                                                                                                                                                                                                                                                                 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                  162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                          GlySerValValIleValGlyArg1leAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                               SerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                  AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (Desert 10, 1998)

Malcolm, B.A., Tareml, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4 cofactor peptide
Patent: US 6211338-A 106 03-APR-2001,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                651
166
13
3
                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 106 from patent US 6211338.
ARI45265
protease and NS4A cofactor peptide
Patent: US 6211338-A 94 03-APR-2001;
Location/Qualifiers
                                                                                                                           Gaps:
                                                                                                                                            US-09-965-594-18 (1-197) x AR145253 (1-651)
                                  /organism="unknown"
188 c 199 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AR145265.1 GI:15107132
                                                                                2.58e-65
870.50
92.82%
85.13%
85.59%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unclassified.
                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
                                               Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unknown
                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                            364
                                                                                                                                                                                                                                   42
                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                         304
                                                                                                                                                                                                                                                                                                                                           102
                                                                                                                                                                                                                                                                                                                                                                               122
                                                                                                                                                                                                                                                                                                                                                                                                 424
                                                                                                                                                                                                                                                                                                                                                                                                                  142
                                                                                                                                                                                                                                                                                                                                                                                                                                   484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR145265
LOCUS
DEFINITION
                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                                 Pred. No.:
         JOURNAL
FEATURES
                                              BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
                                                       ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                               ò
                                                                                                                                                                               g
                                                                                                                                                                                                ò
                                                                                                                                                                                                              g
                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                             243
                                                                                                                                                                                                                                                                                                          244 rGCGTCAACGGCGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303
                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                          SerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer ::::|||||||||
                                                                                                                                                                                                                                               AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998
166
15
12
3
                                                                                          Conservative:
Mismatches:
Indels:
                                                                      Length:
Matches:
                               425
                                                                                                                                           US-09-965-594-18 (1-197) x AR145265 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        completed: August 31, 2003, 00:46:23
He : 2569.57 secs
Location/Qualifiers
1. .1998
/organism="unknown"
a 597 c 567 g
                                                                     9.53e-65
870.50
92.35%
84.69%
85.59%
                                                                                                    Best Local Similarity:
                               409 a
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OW protein - nucleic search, using frame_plus_p2n model Run on: August 30, 2003, 19:13:57 Search time 182.939 Seconds Perfect score: 1019-965-594-18 Perfect score: 1019-965-594-18 Perfect score: 1019-965-594-18 Secoring table: BLOSUM62 Agapop 10.0, Agapext 0.5 Agapop 10.0, Agapext 0.5 Agapop 10.0, Agapext 0.5 Agapop 10.0, Delext 7.0 Searchted: 2552756 seqs, 1349719017 residues Total number of hits satisfying chosen parameters: 5105512 Minimum DB seq length: 0 Maximum Match 10* Post-processing: Minimum Match 10* Maximum Match 10* Post-processing: Minimum Match 10* Listing first 45 summaries Command line parameters: 5105512 Munimum DB seq length: 0 Post-processing: Minimum Match 10* Listing first 45 summaries Listing first 45 summaries Listing first 45 summaries Listing first 45 summaries Command line parameters: 5105512 Maximum Match 10* Post-processing: Minimum Watch 10* Maximum Match 10* Listing first 45 summaries Command line parameters: 51050000000 Post-processing: Minimum Watch 10* Listing first 45 summaries Listing first 5000000000 Loopazir-o DMINES-bits 50000000000 Loopazir-o DMINES-bits 50000000000 Loopazir-o DMINES-bits 5000000000000000000000000000000000000
August 30, 2003, 19:13:57; Search time 182.939 Seconds score: 1037 105-09-965-594-18 107 108-09-965-594-18 108-09-965-594-1
score: 1017 e: IMKKGSVVIGGINLSGDTAVAKAVDFIPVESLETTHRSP 197 table: BLOSUNG2
table: BLOSUM62 Ygapop 10.0, Ygapext 0.5 Ygapop 10.0, Ygapext 0.6 Z552756 seqs, 1349719017 residues 1252756 seqs, 1349719017 residues BB seq length: 0 DB seq length: 2000000000 DB seq length: 2000000000 DB seq length: 2000000000 Cocessing: Minimum Match 100% Maximum Match 100% Listing first 45 summaries Listing first 45 summaries Listing first 45 summaries Tisting first 45 summaries Tisting first 45 summaries Frame
d: 2552756 seqs, 1349719017 residues DB seq length: 0 DB seq length: 0 DB seq length: 0 DB seq length: 10 DB seq length: 0 DB seq length: 100000000 CCESSING: Minimum Match 0% Maximum Match 100% Listing first 45 summaries Listing first 45 summar
unber of hits satisfying chosen parameters: 5105512 DB seq length: 0 DB seq length: 0 DB seq length: 0 DB seq length: 0 BB seq length: 0 BB seq length: 0 BB seq length: 0 BB seq length: 0 Listing first 45 summaries L
DB seq length: 0 DB seq length: 0 DB seq length: 2000000000 DB seq length: 2000000000 Listing first 45 summaries line parameters: Listing first 45 summaries Listing first 45 summaries line parameters: Listing first 45 summaries Listing first 45
processing: Minimum Match 0% Listing first 45 summaries Geneseq_1930n03 -0FWT-fastap -SUFFIX-Ing -WINMATCH-0.1 -LOOPCL-0 Geneseq_1930n03 -0FWT-fastap -SUFFIX-Ing -WINMATCH-0.1 -LOOPCL-0 -45 -DOCALIGN-200 -THR_SCORE-CT -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -45 -DOCALIGN-200 -THR_SCORE-CT -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -45 -DOCALIGN-200 -THR_SCORE-CT -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -45 -DOCALIGN-200 -THR_SCORE-CT -WAIT -DSPBLOCK-100 -LONGLOG WAP -LARGEOURR: WRG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG ITHEOUT-120 -WARN TIMEOUT-30 -THRRADS-1 -XGAPDF-1 -XGAPEXT-0.5 -FGAPOP-6 EXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7 1
nd line parameters: Lframe+_p2n.model -DEF-x1p EXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-biosum62 -TRANS-human40.cdi EXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-biosum62 -TRANS-human40.cdi LLOCAL -OUTFMT—pto -NORM-ext -HRABSIZE-500 -MILEN-0 -ALIGN-15 LLOCAL -OUTFMT—pto -NORM-ext -HRABSIZE-500 -MILEN-0 -MAXIEN-2000000000000000000000000000000000000

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

		æ			SOMETITES	
Result No.	Score	Ouery	Length	DB	qi	Description
		1	J 1	:	-	
. (1017	100.0	594	21	AAA73332	epatitis C v
71	1005	86	200	57	AAA73333	titls C viru
e	1002	. 68	200	7	AAA73331	patitis C viru
+ V	000	. 4	400	7,5	AAA/3334	titis C viru
n ve	0.00	9 6	יי מי מי	7,5	AAA/3350 AAA7330	citis C viru
۰,	976		200	;;	AAA73335	> F ンで
- σο	912	89	588	23	AAA73328	ים מים
6	892.5	87.	~	24	ABA95615	BVDV
10	881.5	86.	1998	70	AAX80355	
11	878.5	86.	σ,	70	AAX80359	4-NS3
12	877.5	98	σ,	50	AAX80353	4-NS3
13	577.5	86.	אית	20	AAX80354	HCV NS4A-NS3 COMPI
.	37.4.5	800	219	572	ABX15706	al sy
ς T •	674.5	96	ρ0	20	AAX80345	NS4A-NS3
7 7	874.5	9 0	סית	2 0	AAX8035/	NS4A-NS3
7 -	074.0	. 0	νс	9 6	AAXBU338	NS4A-NS3
0 0	0.270		ש היים היים	9 6	AAX80352	HCV NS4A-NS3 COMPL
h (0.00		っぃ	2 6	AAA80360	NS4A-NSS
0 7	871.0		Tro	0 0	AAX80349	NS4A-NS3
7 (0.020	, 0 c	100	0 7	AAX80343	HCV NS4A-NS3 compl
77	0.070	0 0	00	7	AAX80344	NS4A-NS3
37	0.070	0	1998	200	AAX80356	NS4A-NS3
4.7	0.00	0	ンソ	7 6	AAXBUSEL	NS4A-NS3
7.0	0.00		0 7	9 0	AAA8U303	NS4A-NS3
9 6	808	0 0	9 0) (AAABUSBS	NS4A-NS3
, o	0.700	000	000	9 0	AAABUS4 /	NS4A-NS3
0 0	2.700		159	9 0	**************************************	NS4A-NS3
67.0	866.7		נאַט	9 0	AAA00331	NOTA NOT
? .	80.00		648	0 0	AAXB0342	NO4A-NOS
32	863.5	8.4	650	0	AAX80346	NSAB-NGS
33	863.5	84	651	20	AAX80350	HCV NS4A-NS3 COMP1
34	861	8		20	AAX23259	mid per-H
35	859	84	σ	50	AAX23258	NS3 DNA.
36	858.5	84.	ø	19	AAV59361	atitis C v
37	858.5	84	Φ	24	ABK87285	cDNA encoding hepa
38	858.5	84.	53	19	AAV59364	itis C
39	858.5	84.	თ	24	ABK87286	titis C v
0.4	858.5	8	9	21	AAZ36212	ectide se
41	854.5	84	m,	10	AAN92097	ined open re
42	854.5	96	സ	07	AAN90327	epatitis C vir
4.	854.5	8	50,	10	AAN92103	ined open rea
4, 4, 4, 1,	80.4 5.74	20 a	7310	0 0	AAN92106 bbwg0336	ed open
•		5	,	4	O C C C C C C C C C C C C C C C C C C C	re nebarir
					ALIGNMENTS	
RESULT]	11					
າ 🗗	73332	standard	1: DNA;	594	G	
AC AAA7	173332;					
19	-DEC-20(00 (flz	irst entr	(<u>k</u>		
DE Hepa'	patitis	C virus	9 NS4A-N	NS3 F	usion protease coding	sequence #5.
KW Hel	epatitis	, NS3 pr lure; li	rotease; iver can	vir cer;	al replication; chronic mutant; mutein; ds.	liver disease;
	Hepatitis Synthetic	C virus	÷			
	;					
ғн кеу	•	-	Location/Qualifier	/ong	lifiers	

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave list replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                   "NS4A-NS3 fusion protein #5"
                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                             Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 26; Fig 15; 66pp; English.
                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                         99US-0115271
                                                                                                                                          06-JAN-2000; 2000WO-US00345
                      / tag a /
                                                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                   P-PSDB; AAB15223
                                                                      WO200040707-A1
                                                                                                                                                                                                                                               Wittekind M,
                                                                                                                                                                             08-JAN-1999;
                                                                                                         13-JUL-2000
```

Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;

	20	40	60	80 240	100	120
Length: 594 Matches: 197 Conservative: 0 Indels: 0 Gaps: 0	MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla 2 	Tyraladinginthrargdlyglugluclycysglngluthrserglnthrdlyargasp 4 	LysasnGlnvalGluGlyGluvalGln11evalSerThralaThrGlnThrPheLeuAla 6 	Thr SerlieAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 8 	SerProlysGlyProvalThrGlnMetTyrThrasnValAsplysAsplæuValGlyTrp 1 	GlnalaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 1
Alignment Scores: Pred. No.: Score: 1017.00 Matches: Percent Similarity: 100.004 Mismatch Query Match: 100.004 Mismatch 100.0	lySerValValIleValC 	hrargGlyGluGluGlyG {{{ }	luGlyGluvalGln1le ⁾ 	lyvalLeutrpthrval: 	roValThrGlnMetTyr' 	lySerArgSerLeuThr
res: 1.3 101 arity: 100 milarity: 100 21 -18 (1-197) x ,	MetLysLysLysG 					GlnAlaProGlnG
Alignment Scores: Pred. No.: Score: Percent Similarity: Best Local Similarity Query Match: DB: US-09-965-594-18 (1-1	Qy 1 Db 1	Qy 21 Db 61	0y 41 0b 121	Oy 61 Db 181	Oy 81 Db 241	Qy 101

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                              301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                          LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                         ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                      CCGCTCGTCACCCTCTTGCTACTTCCGTGCTGTTTCCACCCGTGGTGTTACTAAA
                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis, NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                AlaValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /*tag= a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                          AAA73333 standard; DNA; 594 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0115271.
                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2000-465976,
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13-JUL-2000
                                                                                                                                                     161
                                                                                                                                                                                  481
                                                                                                                                                                                                               181
                                                                                                                                                                                                                                          541
                                                                                          141
                                                                                                                      421
                                                                                                                                                                                                                                                                                                                                        AAA73333;
                                                                                                                                                                                                                                                                            RESULT 2
                                                                                                                                                                                                                                                                                             aaa73333
                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                          g
                                                   셤
                                                                                         ò
                                                                                                                    q
                                                                                                                                                     ò
                             οŽ
```

Page

```
121 LeuValThrArqHisAlaAspValIleProValArqArqGlyAspSerArqGlySer 140
     WO200040707-A1
                              06-JAN-2000;
                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                      41
                                                                                                                                                                                                                                                                                                                                                                                                                                101
                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
480
                                                                                                                                                                                                        120
                                                                                                                                                                                                                                       LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                      240
                                                                                                             09
                                                                                                                                                       80
                                                                                          9
                                                                                                       40
                                                                                                                                                        ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                         GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                              TyralaglnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                              LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                protease; viral replication; chronic liver disease;
liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                          AlaValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         /product= "NS4A-NS3 fusion protein #4"
      other
                        200
44
00
00
                      Length:
Matches:
Conservative:
Mismatches:
Indels:
     0
     Ë
     G; 147
                                                                  (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualiflers
      152
                                                                   US-09-965-594-18 (1-197) x AAA73333
     ပ်
                                                                                                                                                                                                                                                                                                                                                AAA73331 standard; DNA; 594
     A; 191
                       1.8e-85
1005.00
99.49%
98.48%
                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                            1..594
/*tag=
     BP; 104
                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
Synthetic.
                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                       liver failure;
     Sequence 594
                                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000
                  Alignment Scores
                                                                                                                                                                                                                    301
                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                              481
                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                         101
                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                          141
                                                                                                                                                                                                                                                                     421
                                                                                                                                                                                                                                                                                  161
                                                                                                                                                        61
                                                 Query Match:
                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                    q
                                                                                                       ô
                                                                                                                  셤
                                                                                                                                                                             ŏ
                                                                                                                                                                                           g
                                                                                                                                                                                                         ö
                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                        ò
                                                                               ò
                                                                                          g
                                                                                                                                ò
                                                                                                                                        g
                                                                                                                                                       δ
                                                                                                                                                                    g
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3 NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                         Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301 CAGGCTCCGCAGGGTTCCCGTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative:
Mismatches:
Indels:
                                                                                                   Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-18 (1-197) x AAA73331 (1-594)
                                                                                                   Zhang Y,
                                                  င္ပ
                                                                                                                                                                                                                                                                                                                                                       Claim 26; Fig 14; 66pp; English,
                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3.44e-85
1002.00
98.48%
98.48%
                                                                                                   ŝ
99US-0115271
                                                                                                                                                   2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                           P-PSDB; AAB15222
                                                                                                   Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
08-JAN-1999;
```

```
Conservative:
Mismatches:
Indels:
                                  Matches:
                 ength:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73330 standard; DNA; 594 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                995.00
98.98%
97.97%
97.84%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Synthetic.
                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-DEC-2000
   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                  41
                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               481
                                                                                                                                                                                                                                                                                                                                                                                                                                81
                                                                                     Query Match:
                    ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                             đ
                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                 ð
                                                                                                                                                                                                q
                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and hiver cancer. The present invention concerns a number of NS3 mntants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                   480
                                                                                                                  180
                                                                                                                                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /product= "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ပ္ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 26; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                  BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-465976/40.
P-PSDB; AAB15225.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic
                                                                                                                                                                                                                                                                                                                 AAA73334;
                                                                                                                  161
                                                                                                                                                 481
                                                  141
                                                                                     121
                                                                                                                                                                                  181
                                                                                                                                                                                                                  541
                 361
```

Key

RESULT 4 AAA73334

용

οχ

g ŏ Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other;

```
LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                           LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                      180
                                                                                                                                                                                                                                                                         240
                                                                                                                                                                                                                                                                                                 100
                                                                                                                                                                                                                                                                                                                       300
                                                                                                                                                                                                                                                                                                                                              GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                      9
                                                                                                                                                      0
                                                                                                                                                                                                                                                   80
                                                                                                          50
                                                                                                                               9
                                                                                                                                                                                                   1 ATGAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                                                                                                                                                                                                SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                        MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                      TyralaglnglnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; muteln; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag= a
/product= "NS4A-NS3 fusion protein #3"
594
193
000
000
                                                             Gaps:
                                                                                  US-09-965-594-18 (1-197) x AAA73334 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
```

```
(BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 26; Fig 12; 66pp; English.
                                                                                                                                                                                                           BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                           AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2.1e-80
951.00
                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Wittekind M, Weinheimer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      P-PSDB; AAB15220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                      19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                     181
                                                                                                                                                                                                                                        AAA73329;
 141
                               421
                                                                                       481
                                                                                                                                                                                             AAA73329
                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                      õ
                                                                                      a
                                                                                                                     ò
                                                                                                                                                                                                                                                                                            The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the POLYPROTEINS from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 TACGCTCAGCAGACTCGAGGTGAGGAGGGTTGCCAAGAAACCTCCCAGACCGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09
                                                                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21 TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          121 AAAAACCAGGTTGAAAGTGAAAGTTCAGATGGTTTCCACCGCTGCTCAGACCTTCCTGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 ACCIGCATCAACGGIGITIGCIGGACCGITIACCACGGIGCIGGIACCCGIACCATCGCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seguence 594 BP; 103 A; 186 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    594
191
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                     Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (1-594)
                                                                                                                     Zhang Y,
                                                                                                                                                                                                                                                                      Claim 26; Fig 13; 66pp; English.
                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-18 (1-197) x AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.36e-83
985.00
96.95%
96.95%
                                                             99US-0115271
                                                                                                                     Wittekind M, Weinheimer S,
                               06-JAN-2000; 2000WO-US00345
                                                                                                                                                 WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                 P-PSDB; AAB15221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                             08-JAN-1999;
13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
DB:
```

a

ô

g

ò d ó a

ö

ð

a ò 셤 ò

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                           480
                                                                                                                                                      180
                                                                                                                                                                                                                                 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTTGCACCCGTGGTGTTGCTAAA
LeuleuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                      161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                            AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2
                                                                                                                                                                                                                                                                                                                                                 /*tag= a
/product= "NS4A-NS3 fusion protein #2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ocation/Qualifiers
```

us-09-965-594-18.rng

990S-0115271

08-JAN-1999;

9

```
LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LeuLeuSerProArgProIleScrTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                81 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .01 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CIGCIGICCCCGCGTCCGATCTCCTACCTGAAGGTTCCTCCGGTGGTCCGCTGCTGTGC 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                               LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag= a
/product= "NS4A-NS3 fusion protein #8"
     4621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding
     Conservative:
                          Mismatches:
                                                     Indels:
Gaps:
                                                                                                                                 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ВЪ
                                                                                                                                    US-09-965-594-18 (1-197) x AAA73329
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
95.43%
94.92%
93.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
Synthetic.
     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73335
                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAAA73335
XXX
AAA
XXX
AAA
XXX
AAA
XXX
AAA
XXX

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                   ç
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Repatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antivirial treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3*NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 TACGCTCAGCAGACTCGAGGTCTGCTGGGTTGCATCATCACCTCCCTGACCGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                      Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MetLysLysLysGjySerValValileValGlyArgileAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 ATGAAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCCGTACCATCGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TyrAlaGlnGlnThrArqGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArqAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   594
186
0
111
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                      Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-18 (1-197) x AAA73335 (1-594)
                                                                                      Zhang Y,
                                                                                                                                                                                                                                                          Disclosure; Fig 18; 66pp; English.
                                                   ဗ
                                                 (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6.27e-80
946.00
94.42%
94.42%
93.02%
                                                                                      Weinheimer S,
                                                                                                                      2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity:
                                                                                                                                       P-PSDB; AAB15226
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity
Best Local Similar
                                                                                      Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ద
```

140

LeuleuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160

LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer

121 361 141 431

ò

g

ò

300

9

80

~

20 54

541

RESULT 셤

AAA73328;

161

a à

```
21 TyrAlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAsp 40
                                                                                                                                                                                                               101 GluAlaProGluGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                         CCGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                               1 MetLysLysLysGlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pestivirus; Npro; protease; NS3; screening;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
 US-09-965-594-18 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99US-0344456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         04 - DEC - 2001
                                                                                                                                                                                                                                                                            81
                                                                                                                                                                                                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hong 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABA9561
                                                                                           à
                                                                                                                                                                                   a
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                      g
                                                                                                                                                      ò
                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver diseases such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as embling structural studies of the protease and
 180
               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
ProalaglyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
                                                                           Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                            181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 588
182
1
12
2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                 /*tag= a
/product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                         liver failure; liver cancer; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                      BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protease:inhibitor complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9.68e-77
912.00
92.89%
92.39%
89.68%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S,
                                                                                                                                                    AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-465976/40.
P-PSDB; AAB15212.
                                                                                                                                                                                                                                                                                                                      Hepatitis C virus. Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                            WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-JUL-2000
```

Key

120

354

234 100 294

80

174 9

140

414

160

474

180

Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)

Query Match:

œ

us-09-965-594-18.rng

(first entry)

07-SEP-1999

AAX80355;

```
pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
```

Example 2; Columns 17-28; 20pp; English.

The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-peativirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a genome encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a function site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimeric clone of BVDV (bovine viral diarrhea The present sequence is a chimeric clone of BVDV (bovine viral diarrhe virus)/ECV NS3-wt, which was used to illustrate the present invention.

Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;

```
772
                                                                                                                                                                                                                                                                                                                                                                                                                   ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                            832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    413 GGTAGTGTTGTTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472
                                                                                                                                                                                             473 GCCCAGCAGACGAGGCCTCCTAGGGTGTAAGATCACCAGTCTGACTGGCCGGGACAAA 532
                                                                                                                                                                                                                                             592
                                                                                                                                                                                                                                                                                                                                                                      102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GCGGGACACGCCCTGGGCCTATTCAGGGCCGCGGTGTGCCACCCGTGGAGTGGCCAAGGCC 952
                                                                                                                                                                                                                                                                                                                       ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                      GlyServalvalileValGlyArgileAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                       41
                                                                                                                                                                                                                       42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                    773 GTTACGAGGCACGCCGACGTCATTCCCGTGCGCCGGCGAGGTGATAGCAGGGGTAGCCTG
                                                                                                                                                                       22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                               US-09-965-594-18 (1-197) x ABA95615 (1-12734)
       3.06e-73
892.50
92.82%
90.77%
87.76%
                                               Similarity:
                                   Percent Similarity:
Alignment Scores:
                                                                                                                        S
                                                                                                                                                                                                                                                                                                                    82
                                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               953
                                                          Query Match:
DB:
                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                         8
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                      ò
                                                                                                                                                                      à
                                                                                                                                                                                         g
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                     õ
```

standard; cDNA; 1998

AAX80355

RESULT 10

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5 GlySerValValileValGlyArgileAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                       HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998
167
16
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                 CDNA SEQ ID NO:105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-18 (1-197) x AAX80355 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ď,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weber
                                                                                                                    complex encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3.27e-73
881.50
93.37%
85.20%
86.68%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-US2452B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                    HCV NS4A-NS3
                                                                                                                                                                                                                                                                                                                                                                                     W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24 - NOV - 1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                             10-JUN-1999
                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score:
SOCCOCCCC A SEX PRANCE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ą
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 οχ
```

G; 423 T; 0 other;

C; 569

BP; 411 A; 595

Sequence 1998

o,

```
603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NS4A NS3 complex comprising a central hydrophobic domain of native RCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophoble domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the RCV NS3A protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NHR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPRSE activity of NS3. The covalent covalent covalent protease-peptide complexes previously available.
                                                    141
                                                                                                                                                              LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                   ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                              present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                              complex; linker;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linke;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                        ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                     ACV NS4A-NS3 complex encoding cDNA SEQ ID NO:109.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ۲,
ک
                                                                                                                                                                                                                                                                                                                                                        AAX80359 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                           07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                   424
                                                                                                                                                              142
                                                                                                                                                                                                                                                                         182
                                                                                                                                                                                                                                                                                                   604
                                                                                                                                                                                                                                                                                                                                                                                 AX80359
                                                                                                                                                                                                                                                                                                                                       AAX80359
                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                           9 Q
                                                                                                                                                                                                            à
                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                             용
                                                    à
                                                                           g
                                                                                                        ò
                                                                                                                                                                                                                                                                        à
```

```
363
                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                           423
                                                                                                                                                                                                                                                                                                                                                                                                                    141
                                                                                                                                                                                                                                                                                                                                                                                                                                            483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           603
                                                                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                                             303
                                                                                                                                             GGTTCTGTTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                    22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                    122 ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                           244 IGCGTCAACGGCGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                    82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                  102 AlaproGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrileAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                            424 GTCACGAGACATGCTGACGTCATTCCGGTGCGCCGGCGGGGGGCGACAGTAGGGGGGGCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                       GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV: hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:103.
          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                       Gaps:
                                                                                              US-09-965-594-18 (1-197) x AAX80359 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAX80353 standard; cDNA; 1998
         6.26e-73
878.50
93.37%
84.69%
86.38%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          604
                                                                                                                                                                                                                                                                     62
                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAX80353
                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                     à
                                                                                                                                             g
                                                                                                                                                                    ò
                                                                                                                                                                                             g
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                    ò
```

```
22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-1998;
                                                                                                                                                                                                                                                                     07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                       SON COCCOCCOCCON TO THE SERVING SON THE SERVIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                124 TCCCAACAGGGGGGGGCCTACTTGGTTGCAAGATCACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            81
                                                                                                                                                                                                                                                                                                                              NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 competed. and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active then non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                            present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1998
167
15
11
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
Gaps:
                                                                                                                                               Yao N;
                                                                                                                                                                                                                              New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                       Disclosure; Page 160-162; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-18 (1-197) x AAX80353 (1-1998)
                                                                                                                                               Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.77e-73
877.50
92.86%
85.20%
                                       98US-0094331
97US-0067315
98WO-US24528
                                                                                                                                             Taremi SS,
                                                                                                        (SCHE ) SCHERING CORP
                                                                                                                                                                                  WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                           28-JUL-1998;
28-NOV-1997;
24-NOV-1998;
                                                                                                                                               Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             :.
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex domain. Where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arbase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4a cofactor; NS4a-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                 182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                            Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1998
167
15
11
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 163-166; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JS-09-965-594-18 (1-197) x AAX80354 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Weber PC,
                                                                                                                                                                                                                                                           AAX80354 standard; cDNA; 1998 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7.77e-73
877.50
92.86%
85.20%
86.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity:
```

```
483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    243
                                                                                                                                                                                     LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           162 AlaGlyHisAlaValGlyIlePheArqAlaAlaValSerThrArqGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                has
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C; ds; viral prototoxophore; anti-viral; tumour; virus; infection; antitumour; toxophore; human immunodeficiency virus; HIV infection; herpes simplex virus; HSV; rhinovirus; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGTGGTCCACTGCTCTGCCCCT
                                                                                Novel synthetic viral prototoxophore for treating viral infections, toxin molety incorporated into substrate domain specific for viral enzyme, bound and modified by viral enzyme to get converted into toxophore
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Anti-viral synthetic prototoxophore associated DNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Shepard
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 1; Page 62; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Neuteboom STC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26-APR-2002; 2002WO-US13223.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABX15706 standard; DNA; 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27-APR-2001; 2001US-286893P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-MAR-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (NEWB-) NEWBIOTICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2003-167102/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200287500-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Cathers BE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-NOV-2002
                                                                                   42
                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                                                                        364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABX15706;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ABX15706

IID ABX1

XX

XX

XX

XX

XX

ABX1

XX

ABX1

XX

ABX1

BE

ABX1

BE

ABX2

BE

ABX3

ABX3

ABX4

ABY4

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ă
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                             셤
                                                                                                                                                                       ç
                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                à
```

```
comprising a toxin moiety operatively incorporated into a substrate

domain specific for a viral enzyme. This prototoxophore may be bound
and anddified by the viral enzyme. This prototoxophore may be bound
and anddified by the viral enzyme thus converting it to a toxophore.

Condition of the invention is a method for enhancing the anti-viral
effect of an antiviral agent, this method comprises contacting a cell,
infected with a virus or is susceptible to infection, with a
prototoxophore. The invention further comprises an assay to identify
anti-viral agents, comprising contacting an infected cell with a
candidate agent and comparing the ability of the agent to inhibit the
conforting way have virusing the ability of the agent to inhibit the
prototoxophores of the invention may be useful for reducing or
infected with a virus of is susceptible to infection with a virus, with
an effective amount of the prototoxophore. The cells are cell lines
adapted to long term continuous culture or isolated from a subject.
The prototoxophore is also useful for amellorating the severity of a
viral infection in a subject, where the virus is selected from human
immunodeficiency virus (HIV), herpes simplex virus (HSV), rhinovirus and
hepatitis virus, by administering an effective amount of the
prototoxophore to the subject. The prototoxophores of the invention are
also useful for treating tumours. The present sequence represents an
antiviral prototoxophore associated DNA sequence, this sequence is
the invention although in the colour prototon protote in a prototox
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GGTAGTGTGGTCATTGTGGGTAGGATCATTTTGTCCGGTAGTGGTAGTATCACGGCGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      102 AlaproGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GTCACGAGGCACGCCGATGCTTCCCGTGCGCGGGGGGGTGATAGCAGGGGCAGCCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrlleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the invention although it is clearly not a protein sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 612 BP; 120 A; 171 C; 191 G; 130 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (1-612)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-18 (1-197) x ABX15706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3.39e-73
874.50
91.28%
89.74%
85.99%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
        $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  òχ
```

AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181

162

This invention relates to a novel synthetic viral prototoxophore

g

```
162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                            124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCTTACAGGCCGGGACAAG 183
                                                                                                                                  62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr1leAlaSer
                           AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                             182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                           42
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                g
                                               合
                                                                               ò
                                                                                                     셤
                                                                                                                                    à
                                                                                                                                                               셤
                                                                                                                                                                                           ò
                                                                                                                                                                                                                a
                                                                                                                                                                                                                                               ő
                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present Sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NWR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPses activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
 499 GCGGGGCACGCCGTGGCCATATTTAGGGCCGCGGTGTGCACCCGTGGAGTGGCTAAGGCG 558
                                                                                                                                                                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                          ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                        559 GIGGACTITATCCCTGTGGAGAACCTAGAGACAACCATGAGGTCC 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    651
166
10
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0 other;
                                                                                                                                                                                            HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 147-148; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weber PC,
                                                                                                            BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3.66e-73
874.50
93.33%
85.13%
85.99%
                                                                                                           AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                                                98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                       98WO-US24528
                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-385385/32
                                                                                                                                                                                                                                                                           Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                   W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                  07-SEP-1999
                                                                                                                                                                                                                                                                                                                                             10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Malcolm BA,
                             182
                                                                                                                                      AAX80345;
                                                                                                              셤
```

윱 οŽ

US-09-965-594-18 (1-197) x AAX80345 (1-651)

Conservative: Mismatches: Indels:

Percent Similarity: Best Local Similarity:

Query Match: DB:

AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSerGlnThrGlyArgAspLys 41

```
648
completed: August 30, 2003, 19:48:09 e : 188.939 secs
```

```
GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.
```

```
August 30, 2003, 19:20:43; Search time 1910.31 Seconds (without alignments) 2506.388 Million cell updates/sec
                                                                                                                                                                 1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
OM protein – nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                       BLOSUM62
Xgapop 10.0 , Xgapext C
Ygapop 10.0 , Ygapext C
Fgapop 6.0 , Fgapext 7
Delop 6.0 , Delext 7
                                                                                                                         US-09-965-594-18
1017
                                                                                                                         Title:
Perfect score:
                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                   Sequence:
                                        Run on:
```

45562784 22781392 seqs, 12152238056 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Command line parameters:
-WODEL-frame+_p2n.model -DEV-xlp
-Q-Qqn2_1/USPTO_spool/US0995594/runat_29082003_151919_28322/app_query.fasta_1.2872
-Q-Cqn2_1/USPTO_spool/US09965594/runat_29082003_151919_28322/app_query.fasta_1.2872
-Q-Cqn2_1/USPTO_spool/US09965594/runat_29082003_151919_28322/app_query.fasta_1.2872
-UNTS-b1ts -GTART-1 -EMD-1 -MATRIX-blosum62 -TRANS-human40.cd1 -11.ST-45
-UDCALIGN-200 -THR_SCORE-pct -THR_MAX_100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL
-UDCALIGN-200 -THR_SCORE-pct -THR_MAX_100 -THR_MIN-0 -MAX_LEN-200000000
-USER-US0995594_GCM_1_1.18630_Grunat_29082003_151919_28322 -NCPU-6 -ICPU-3
-NO_MMAP -LANGEQUERY -NEG_SCORES_0 - WAIT -DSPBLOCK-100 -LONGLOG
-DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THRENDS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6

em_gss_pln:* em_gss_vrt:* em_gss_fun:* gb_est2:*
gb_htc:*
gb_est3:*
gb_est4:*
gb_est5:*
em_estfun:* em_gss_mus:* em_estom:* em_gss_hum:* em_gss_inv:* em_gss_mam:* em_estba:*
em_esthum:*
em_estin:*
em_estuu:*
em_estov:* em_estro:* em_htc:* gb_estl:* EST:* Database :

en_gss_pro: en_gss_rod: en_gss_phg: en_gss_vrl: en_gss_vrl: en_gss_vrl: en_gss_rri: en_gs_rri: e

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Ouery Court Cour			•					
106 10.4 1146 12 12 12 12 12 13 14 12 12 12 13 14 12 12 12 13 14 12 12 12 13 13 14 12 12 12 13 13 14 12 12 12 13 13 14 12 12 12 13 13 14 12 12 12 13 13 14 12 12 12 12 13 13 14 12 12 12 13 13 12 12 12	Sult No.	Scor	Query Match	Length	90	13	scripti	
2 104.5 10.3 1031 14 CB950999 CB95 2 102.5 10.1 1403 13 BQ952401 BQ952401 BQ95.5 9.4 502 9 AA036844 BP304699 BQ5.5 9.4 701 10 BF663444 BP304699 BQ5.5 9.4 984 10 BF630464 BP304699 BQ5.5 9.4 984 10 BF630464 BP304699 BQ5.5 9.3 9.1 1213 12 BM553374 BM560704 CC406704 BM560704 CC406704 BM560704 BM560707 BM760704 BM560704 BM560704 BM560704 BM560704 BM560704 BM560707 B	•	10	10.4	1146	12	BM91580	915803	: ×
3 102.5 10.1 1403 13 BO926101	. 4	104.	ö	03	14	66056	6609	AGENCOURT
4 100.5 9.9 1199 13 BAG892447 BAG884	(*)	102.	ö	0	13	92610	929	C)
95.5 9.4 701 9 And 36834 Abods 75.5 9.4 701 10 BR663444 BP304699 8 95.5 9.4 701 10 BF304699 BF30469 93.5 9.2 962 10 BF304699 BF30699 BF30699 93.5 9.2 962 10 BF304659 BF306 BF304 93 9.1 1213 12 BM553374 BF806 BF806 1 9.1 128 12 BM402866 BM40 BM5028642 BM408 1 701 29 BF34288 BF806 BK4177 BM508 BK4177 BM408 BK418	4 (100.	•	13	13	39248	2892487	8
95.5 9.4 701 10 BFR63344 BFR96 95.5 9.4 701 10 BFR63344 BFR96 95.5 9.4 994 10 BFR96499 BFR96 95.5 9.2 9.3 629 10 BG089727 BG08 93.5 9.2 1733 12 BM553374 BM55 93.5 9.1 1213 12 BM553374 BM55 92.5 9.1 772 29 BX23898 BBR24 BM55 92.5 9.1 772 29 CC406704 CC40 92.5 9.1 789 13 BM2653399 BM76 92.5 9.1 789 866 13 BM219343 BM2193 BM219343 BM219333463 BM219333463 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM2193333463 BM2193 BM2193 BM219333463 BM2193 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM2193 BM219333463 BM2193 BM219333463 BM2193 BM219333463 BM219333463 BM219333463 BM219333463 BM219333463 BM219333463 BM219333463 BM2193 BM219333463 BM2193333463 BM219333463 BM219333463 BM219333463 BM219333463 BM219333463 BM219333443 BM21933344	٠, ١	95.		0	o i	36834	036834	k29d05.r
95.5 9.4 984 10 BF304699 BF789 93.5 9.2 90.6 10 BF304699 BF789 93.5 9.2 90.5 13 BM542842 BM759 93.5 9.2 90.5 13 BM542842 BM759 93.5 9.1 528 12 BM540266 BM759 93.5 9.1 1213 13 BM541277 BM523334 BM759 92.5 9.1 172 29 CC406705 CC406705 GC406705 BM759 91.5 9.0 1294 13 BM541277 BM5289 BM759 91.5 9.0 1294 13 BM541377 BM759 91.5 9.0 1294 13 BM761998 BM759 91.6 90.5 80.9 1440 12 BM76199 BM759 90.5 80.9 1440 12 BM76199 BM759 BM759 90.5 80.9 1141 11 AK080645 BM759 BM759 90.6 80.8 653 12 BM855317 BM861059 BM759 90.8 8 652 12 BM855317 BM861059 BM759 90.8 8 652 12 BM855317 BM861059 BM759 90.8 8 8 1015 13 BM52851 BM759 90.8 8 8 1015 13 BM357770 BM759 90.8 8 8 1015 13 BM357770 BM759 90.8 8 8 1015 13 BM35849 BM759 BM759 90.8 8 8 1015 13 BM35849 BM759 90.8 8 8 1015 13 BM35849 BM759 90.8 8 8 1015 13 BM338469 BM339469 BM	ا ب	95.		~ I	2	36324	324	96304200
95 9.3 629 10 BG089727 BG08 93.5 9.2 965 13 BU542842 BW55 93.5 9.2 1733 12 BM4523848 BW45 93.9 9.1 644 29 BX238988 BW45 93.9 9.1 701 29 BX238988 BW45 93.9 9.1 1213 13 BU541777 CC40 92.5 9.1 772 29 CC406704 CC40 91.5 9.0 1294 13 BQ894657 BQ92 91.5 9.0 1294 13 BQ895657 BQ92 91.5 9.0 1294 13 BQ252649 AQ62 91.5 9.0 1294 14 CA728999 BG88 90.5 8.9 144 11 CA728999 BG88 90.5 8.9 144 11 AX080479 BW70 90 8.8 622 9 AV85401 BW46 90 8.8 623 1 BW565338 BW46 90 8.8 623 1 BW5653 90 8.8 623 1 BW5653 90 8.8 623 1 BW361659 BW46 90 8.8 631 10 BW361659 BW46 90 8.8 631 10 BW361659 BW46 90 8.8 631 10 BW361659 BW364 90 8.8 123 13 BX39847 BX39 90 8.8 1043 29 CC406AW BW36 90 8.8 1043 13 BX39849 BX33 90 8.8 1043 29 CC406AW BW36 90 8.8 1043 29 CC406AW BW36 90 8.8 1049 29 CC406AW BW36 90 8.8 1201 3 BX338469 BW36 90 8.8 1049 29 CC406AW BW36 90 8.8 1049 20 CC406AW BW36 90 80 80 80 80 80 80 80 80 80 80 80 80 80	-	95.	٠	т	10	30469	9469	60188825
9 93.5 9.2 9.2 173 12 BM553374 BM55 9 3 9.1 528 12 BM553374 BM55 9 3 9.1 644 29 BX238988 BX23 9 3 9.1 701 29 BX238988 BX23 9 3 9.1 702 29 BX238988 BX23 9 2.5 9.1 772 29 CC406704 CC40 9 2.5 9.1 772 29 CC406704 CC40 9 2.5 9.1 789 29 CC406705 CC40 9 1.5 9.0 1294 13 BQ894657 BQ894 9 1.5 9.0 1294 13 BQ8926457 BQ899 9 1 8.9 66 13 BQ25457 BQ899 9 1 8.9 866 13 BX43207 BX43 9 0 8 8 653 12 BM625319 BU70 9 0 8 8 653 12 BM625317 BW70 9 0 8 8 653 13 BX43207 BX43 9 0 8 8 653 13 BX22565 BX33 9 0 8 8 1049 29 CX80479 BX33 9 0 8 8 1049 29 CX80404W BX26 9 0 8 8 105 13 BX33847 BX33849 BX33 9 0 8 8 1049 29 CX80404W BX26 9 0 8 8 1049 29 CX80404W BX26 9 0 8 8 1049 29 CX80404W BX38 9 0 8 8 105 13 BX338469 BX33 9 0 8 8 1201 3 BX338469 BX33 9 0 8 8 1201 3 BX338469 BX33 9 0 8 8 1201 3 BX338469 BX33	ω.	σ,	•	CV.	10	38972	39727	mab90e06.
0. 93.5 9.2 1733 12 BM553374 BM55 93 9.1 528 12 BM402566 BM23 93 9.1 701 29 BZ34281 BZ34 92.5 9.1 772 29 CC406704 CC40 92.5 9.1 772 29 CC406705 CC40 92.5 9.1 772 29 CC406705 CC40 91.5 9.0 1294 13 BQ52457 BQ92 91.5 9.0 1294 13 BQ52457 BQ92 91.5 9.0 1294 13 BQ52457 BQ92 91.5 9.0 1294 13 BQ52457 BQ92 91.8 9 66 13 BQ52457 BQ92 91.8 9 66 13 BQ25457 BQ43 90.5 8.9 1440 12 BQ52399 BG88 90.5 8.9 1440 12 BM467291 BX43 90.6 8.8 659 12 BM467299 BM70 90 8.8 653 13 BM291059 BM83 90 8.8 123 13 BM291059 BM83 90 8.8 123 13 BM291059 BM83 90 8.8 1043 329469 BM33 90 8.8 1043 339469 BM33 90 8.8 1049 29 CM5040AM BM23 90 8.8 1201 3 BX338469 BM33 90 8.8 1201 3 BX338469 BM33 90 8.8 1201 3 BX338469 BM33 90 8.8 1201 3 BX338469 BM33	o,	93.		90	13	54284	12842	AGENCOURT
93 9.1 528 12 BM402566 BN428 93 9.1 644 29 BX238988 BN428 93 9.1 701 29 BX238988 BN428 93 9.1 1213 13 BU54177 92.5 9.1 772 29 CC406704 92.5 9.1 772 29 CC406704 92.5 9.1 789 29 CC406704 91.5 9.0 1294 13 BQ92547 91 8.9 646 13 BQ92547 90 8.8 651 13 BM402319 90 8.8 652 9 1440 12 BM402319 90 8.8 631 10 BM6231 90 8.8 653 13 BM21943 90 8.8 653 13 BM21943 90 8.8 653 13 BM2517 90 8.8 653 13 BM2518 90 8.8 653 13 BM258 90 8.8 653 13 BM258 90 8.8 123 13 BM25256 90 8.8 123 13 BM35471 90 8.8 123 13 BM35471 90 8.8 1043 29 CM338 90 8.8 1049 29 CM348 90 8.8 1040	ä	93.		73	12	55337	33374	AGENCOURT
93 9.1 644 29 BX238988 BX238 94 9.1 701 29 BZ3423811 BZ344 95 9.1 1213 13 BG541277 CC40 92.5 9.1 772 29 CC406704 CC40 91.5 9.0 528 28 A062049 BQ84657 91.5 9.0 1294 13 BQ925457 BQ92 91.5 9.0 1294 13 BQ925457 BQ92 91.8 9.0 646 13 BQ25457 BQ92 91.8 9.0 646 13 BQ253399 BG740 91.8 9 966 13 BX44207 BX43 90.5 8.9 1141 11 AX080349 BG740 90.6 8.8 659 12 BM625317 BX43 90.8 8.8 651 10 AW961059 BW70 90.8 8.8 653 13 BX35770 BX43 90.8 8.8 653 13 BX35770 BX35 90.8 8.8 104 91 BX35770 BX35 90.8 8.8 104 91 BX35770 BX35 90.8 8.8 104 91 BX35770 BX35 90.8 8.8 104 92 CX080479 BW70 90.8 8.8 653 13 BX35770 BX35 90.8 8.8 104 92 CX080404 BX33 90.8 8.8 120 13 BX338469 BX33 90.8 8.8 120 13 BX338469 BX33 90.8 8.8 120 13 BX338469 BX33	=	6		\sim	12	10256	12566	SLACOSF12
93 9.1 701 29 B2342381 B234 92.5 9.1 772 29 CC406704 CC40 92.5 9.1 772 29 CC406705 CC40 92.5 9.1 772 29 CC406705 CC40 91.5 9.0 528 28 AC620049 AC620 91.5 9.0 1294 13 BQ924657 AQ620 91.5 9.0 1294 13 BQ925457 AQ620 91.8 9 64 12 BQ925457 BQ92 91.8 9 64 12 BQ925457 BQ92 91.8 9 966 13 BQ42913 BQ420 90.5 8 9 1440 11 BQ92543 BQ20 90.5 8 9 1440 11 BW7080045 BW70 90 8 8 651 10 BW3040 BW70 90 8 8 651 10 BW3040 BW304 90 8 8 653 13 BC22566 BW80 90 8 8 653 13 BC22566 BW30 90 8 8 8 753 13 BC22566 BW30 90 8 8 8 12 13 BC22566 BW30 90 8 8 1015 13 BX39463 BX33 90 8 8 1049 29 CC8040AW AL562 90 8 8 1201 3 BX338469 BX33 90 8 8 1201 3 BX338469 BX33	7	5		-	29	3898	88688	Danio rer
93 9.1 1213 13 BU541777 CC40 92.5 9.1 772 29 CC406704 CC40 6704 CC40 6705 92.5 9.1 772 29 CC406704 CC40 6705 92.5 9.1 789 29 CC406704 CC40 6705 92.5 9.1 789 29 CC406705 GC40 91.5 9.0 1294 13 BQ894657 BQ8926474 91.5 9.0 1294 13 BQ8926547 CA728998 CA728998 646 12 BQ8026547 CA728999 BQ8026547 CA728999 BQ90 BG6 13 BU402910 BU40 BU40 90.5 BQ90 BG6 13 BU402910 BU40 BU40 90.5 BQ90 BG6 13 BU40945 BW809 90 BG8 659 12 BW802517 BW809 90 BG8 659 12 BW802517 BW809 90 BG8 659 12 BW802517 BW809 90 BG8 659 12 BW805517 BW809 90 BG8 659 13 BX903647 BW809 90 BG8 1123 BX903647 BX309 BX309 BW309 BW90 BW90 BW90 BW90 BW90 BW90 BW90 BW	Ξ	σ		0	29	1238	12381	1c83b11.b
92.5 9.1 772 29 CC406704 CC40 92.5 9.1 772 29 CC406705 GC40 92.5 9.1 772 29 CC406705 GC40 92.5 9.1 789 29 CC406705 GC40 92.5 9.1 938 13 BQ894657 BC896 91.5 9.0 1294 13 BQ25457 BQ892 91.5 9.0 1294 13 BQ25457 BQ892 91.5 9.0 1294 13 BQ25457 BQ892 91 8.9 866 13 BQ25499 BQ40 91 8.9 866 13 BQ40243 BQ402 90.5 8.9 1141 11 ANO8034 BQ402 90.5 8.9 1141 11 ANO8034 BAC02 90 8.8 569 12 BM45217 BW202 90 8.8 653 12 BM452317 BW202 90 8.8 653 12 BM825317 BW202 90 8.8 653 13 BX95492 BW202 90 8.8 757 12 BM825317 BW202 90 8.8 757 12 BM825317 BW202 90 8.8 1049 29 CX80404 BW202 90 8.8 1049 8.8	7	6		2	13	1177	11777	AGENCOURT
92.5 9.1 789 29 CC406705 CC40 92.5 9.1 789 29 CC406705 91.5 9.0 1294 13 BQ894657 A0620049 91.5 9.0 1294 13 BQ894657 A0620049 91.5 9.0 1294 13 BQ826457 A0620049 91.5 9.0 1294 13 BQ826457 BQ898 BG89 91.8 99 646 12 BG85399 BG85399 BG85399 BG85399 BG85399 BG89 BG85399 BG89 BG89 BG89 BG89 BG89 BG89 BG99 BG9	15	92.		~	59	0690	5704	PUHKD12TB
92.5 9.1 938 13 BQ894557 BQ89 9 1.5 9.0 1294 13 BQ894557 MQ624 9 1.5 9.0 1294 13 BQ925547 MQ629 9 1.5 9.0 1294 13 BQ925547 MQ629 9 1.8 9.9 646 12 BQ825547 MQ629 9 1.8 9.9 646 12 BQ8299 BQ829 BQ929 B	16	92.		8	5	0670	5705	PUHKD12TC
91.5 9.0 528 28 AQ62049 AQ6299 91.5 9.0 1294 13 BQ925457 BQ929 91.5 9.0 1294 13 BQ925457 BQ929 91.5 9.0 1294 13 BQ925457 BQ929 91.8 9.9 646 12 BG83399 BG83399 BG83399 BG83399 BG83399 BG929 BG83399 BG929 B	ñ	92.		~	13	9465	1657	AGENCOURT
91.5 9.0 1294 13 BQ925457 BQ92 91.5 9.0 1294 13 BQ925457 BQ92 91 8.9 646 12 BG853999 BG85399 BG8599 BG8999 BG9999 BG99999 BG99999 BG9999 BG9999 BG9999 BG99999 BG99999 BG99999 BG99999 BG9999 BG99999 BG99999 BG99999 BG99999 BG99999 BG99999 BG99999 BG9999 BG99999 BG99999 BG99999 BG99999 BG99999 BG99999 BG99999 BG9999 BG99999	ä	91.		~	28	2024	1249	HS_5182_E
0 91 8.9 580 14 CA728998 CA728998 1 91 8.9 646 12 BG885 BG885 3 91 8.9 966 13 BU4021943 BU4021943 4 90.5 8.9 966 13 BU4021943 BU4021943 5 90.5 8.9 914 11 AK080479 BX403 6 90.5 8.9 1141 11 AK080479 BX403 9 8.8 50 12 BM408077 BM408 9 8.8 50 12 BM82517 BX408 9 8.8 617 10 BK910 BX408 9 8.8 621 10 BX835401 AW956159 AW956159 9 8.8 631 10 BW8647 BX409 BX409 9 8.8 631 10 BW8647 BX405 BX405 9 8.8	5	91.		29	13	2545	5457	AGENCOURT
91 8.9 646 12 BGB53399 BGB5 91 8.9 753 13 BUJ02910 BUJ0 91 8.9 966 13 BUJ1943 BUJ1944 90.5 8.9 1141 11 AKO80645 BAX08 90.5 8.9 1140 11 BMC080779 BAX08 90 8.8 569 12 BMG52317 BMG2 90 8.8 651 12 BMG55318 BMG2 90 8.8 652 12 BMB25317 BMG2 90 8.8 652 12 BMB25317 BMG2 90 8.8 653 12 BMG2531 BMG3 90 8.8 653 12 BMG2531 BMG3 90 8.8 757 12 BMS2540 BMX3 90 8.8 1015 13 BX39459 BX33 90 8.8 1049 29 CXS040AW ALS6 90 8.8 1020 13 BX33463 BX33 90 8.8 1049 29 CXS040AW ALS6 90 8.8 1201 3 BX333463 BX33463 BX333 90 8.8 1201 3 BX333463 BX333463 BX333463 BX333 90 8.8 1201 3 BX333463 BX333463 BX333463 BX333	2	6		TO)	14	339	339	
2 91 8.9 753 13 BU4402110 BU440 3 9.9 866 13 BU42193433 BU421 5 90.5 8.9 814 11 ANS09179 BX434207 6 90.5 8.9 814 11 ANS09179 BX40 8 9.0 8.8 9.0 1440 12 BM467279 BN468 9 9.0 8.8 50 12 BM708007 BM408 BM408 9 9.0 8.8 617 10 BM52317 BM508 BM508 10 9.0 8.8 617 10 AW961059 BM978 BM978 10 9.0 8.8 622 9 AW830640 BW8961059 AW896 10 8.8 631 10 AW830640 BW8961059 BW8961059 <td>5</td> <td>o</td> <td></td> <td>•</td> <td>12</td> <td>399</td> <td>399</td> <td></td>	5	o		•	12	399	399	
91 8.9 866 13 BUZ19343 BUZ1944 4 91 8.9 906 13 BX4207 BX4207 5 90.5 8.9 1141 11 AK080045 AK08 8 90 1141 11 AK080045 BWA6 9 0.6 8.8 569 12 BM25317 BWA6 9 0 8.8 617 10 BW35401 BWA9 9 0 8.8 622 9 AV835401 BW39 9 0 8.8 631 10 AW961059 AW96 9 0 8.8 631 10 AW961059 BW39 9 0 8.8 875 12 BM39847 BW39 9 0 8.8 875 13 BX29864 BW39 9 0 8 8 812 13 BX29854 BY39 9 0 8 8 812 13 BX39847 BW39 9 0 8 8 812 13 BX39847 BW39 9 0 8 8 812 13 BX39847 BY39 9 0 8 8 812 13 BX39849 BX39 9 0 8 8 1043 4921 BX39 9 0 8 8 1049 29 CXS040AW AL562 1 0 0 8 8 1201 3 BX39849 BX33 1 0 0 8 8 1201 3 BX39849 BX39 1 0 0 8 8 1201 3 BX339469 BX39 1 0 0 8 8 1201 3 BX339469 BX39 1 0 0 8 8 1201 3 BX339469 BX33	2	σ		ın	13	291	391	
4 91 8.9 906 13 BX44207 BX43 6 90.5 8.9 814 11 AK080479 AK08 7 90.5 8.9 1141 11 AK080479 AK08 7 90.5 8.9 1141 11 BM62779 BX408 90 8.8 50 12 BM625317 BX408 BX408 90 8.8 67 10 BX625318 BX608 BX608 1 90 8.8 631 10 BX6101 AW96 3 90 8.8 631 10 AW96 AW96 4 90 8.8 753 12 BX128 BX13 BX03 5 90 8.8 812 13 BX03 BX03 BX12 BX20 6 90 8.8 812 13 BX20 BX3 BX20 8 90 8.8 104 8.8	2	σ	•	o	13	334	334	
5 90.5 8.9 814 II CNS09179 BXOS 7 90.5 8.9 1141 11 MAC080545 MAC08 9 90.5 8.9 1440 12 BM467279 BXAC08 BM408 BM	7	6	•	0	13	120	5	
6 90.5 8.9 1441 11 AK08 AK08 9 9.6 8.8 50 12 BM467279 BM467779 BM467779 BM467779 BM467779 BM467779 BM467779 BM4708000 BM708000 AV808000 AV808000 AV808000 BM708000 BM7080000 BM7080000 BM7080000<	7	96	•	8	11	317	966	
90.59 9.90 9.80 9.90 9.80 9.80 9.80 9.80 9.8	7	0	•	Ξ:	11	054	2	
8 90 8.8 569 12 BM708007 BM82 9 90 8.8 569 12 BM825517 BM82517 BM82517 BM82517 BM825117 BM82	7	.00	•	4	12	5727	727	
99 90 8.8 569 12 BM825317 BM82 10 90 8.8 617 10 BM855318 BE055 2 90 8.8 613 10 AW85401 AW865 3 90 8.8 631 10 AW861059 AW86 4 90 8.8 658 12 BM830847 BM83 5 90 8.8 763 13 BX093894 BM83 6 90 8.8 812 13 BX093894 BX09 7 90 8.8 812 13 BX093894 BX09 9 90 8.8 1015 13 BX404631 BX40 9 90 8.8 1015 13 BX404631 BX40 9 90 8.8 1123 13 BX39849 AX865 1 90 8.8 1201 13 BX39849 AX565 1 90 8.8 1201 13 BX339469 BX33 1 564 1201 13 BX339469 BX33 1 564 12 BG827880 BX33	~	6	•	0	12	0000	800	
0 90 8.8 617 10 BE055938 BED50 2 90 8.8 622 9 AV835401 AV8354 3 90 8.8 631 10 AW961059 AV835401 4 90 8.8 757 12 BM830847 BLM83 5 90 8.8 757 12 BM830847 BLM83 5 90 8.8 757 12 BM830847 BLM83 6 90 8.8 753 13 BX093694 BX09 6 90 8.8 12 13 BX222616 BX35 9 90 8.8 1015 13 BX30849 AL562 1 90 8.8 1201 13 BX33849 AL563 1 90 8.8 1201 13 BX33849 BX33 1 0 8.8 1201 13 BX338469 BX33	2	6	•	ω.	12	2531	\sim	K-EST0097
1 90 8.8 622 9 AV835401 AV835 2 90 8.8 631 10 AM961059 AN96 4 90 8.8 658 12 BM830847 BM830845 5 90 8.8 757 12 B1258551 B125 5 90 8.8 757 12 B1258551 B125 5 90 8.8 757 12 B125851 B125 9 90 8.8 812 13 B0222616 B022 9 90 8.8 1015 13 B1257770 B125 9 90 8.8 1049 29 CNS040AM AL26 1 90 8.8 1201 9 AL562877 AL562 2 90 8.8 1201 13 B1339469 B125 1 50 8.8 1201 13 B1339469 B125 1 564 12 BG827880 B125 1 50 8.8 1201 13 B1339469	<u>ج</u>	σ	•	-	10	BE05593	9	945020D03
2 90 8.8 631 10 AW961059 AW96 3 90 8.8 658 12 BM830847 BM830 5 90 8.8 753 12 BM830847 BM830 5 90 8.8 753 12 BX093694 BM830 7 90 8.8 812 13 BX093694 BX09 8 812 13 BX093694 BX09 9 90 8.8 812 13 BX0434921 BX0430 9 90 8.8 1015 13 BX39849 BX39 1 90 8.8 1029 BX39849 BX39 2 90 8.8 1201 3 BX39869 BX33 2 90 8.8 1201 3 BX33969 BX33 2 90 8.8 1201 13 BX33969 BX39 2 80 8.8 1201 13 BX33969 BX39 2 90 80 80 80 80 80 80	3	6	•	N	5	V835401	35	AV835401
3 90 8.8 658 12 BMB30847 BMB3 4 90 8.8 757 12 B1258851 BXD3 5 90 8.8 763 13 BQ434921 BXD43 7 90 8.8 812 13 BQ434921 BQ43 8 90 8.8 812 13 BQ222616 BXD3 9 90 8.8 1015 13 BX404631 BX405 1 90 8.8 1049 29 CNS040AW BX30 1 90 8.8 1201 9 AL562877 AL562 3 90 8.8 1201 13 BX339469 BX33 4 6 8.8 1201 13 BX339469 BX33 5 90 8.8 1201 13 BX339469 BX33 5 90 8.8 1201 13 BX339469 BX33 6 80 8.8 1201 13 BX339469 BX33 6 8 8 1201 13 BX339469 BX33	'n	O	•	3	10	5105		EST373026
4 90 8.8 757 12 B12288551 B125 5 90 8.8 763 13 BX093694 BX093694 7 90 8.8 812 13 BX34921 BX093694 BX093694 7 90 8.8 812 13 BX222016 BX022 9 8.8 18 87 12 BX35 9 8.8 1015 13 BX404651 BX45 1 90 8.8 1123 13 BX3849 BX39 1 90 8.8 1201 9 AL562877 BX339 4 8 1201 13 BX339469 BX33 8 9 8 1201 13 BX339469 BX36 8 9 8 1201 13 BX339469 BX36	m	on	٠	S	13	3084		K-EST0104
5 90 8.8 763 13 BX093694 BX09 6 90 8.8 812 13 BQ434921 BQ43 9 90 8.8 812 13 BQ422616 BQ22 8 90 8.8 1015 13 BX357770 BX367 9 90 8.8 1049 29 CNS040AW AL2662 1 90 8.8 1201 3 BX39849 BX39 9 90 8.8 1201 13 BX39869 BX33 5 90 8.8 1201 13 BX33969 BX33 5 90 8.8 1201 13 BX339469 BX33 5 90 8.8 1201 13 BX339469 BX33 7 0 8.8 1201 13 BX339469 BX33	ž	σ		w	12	5885		602969666
6 90 8.8 812 13 BQ434221 B043 9 9 8.8 859 13 BQ222616 BX22 9 90 8.8 1015 13 BX404631 BX40 0 90 8.8 1049 29 CNS040AW AL26 1 90 8.8 1201 3 BX398349 BX39 2 90 8.8 1201 9 AL562877 AL562 3 90 8.8 1201 13 BX39469 BX33 4 90 8.8 1201 13 BX339469 BX33 5 5 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	œ,	6	•	9	13	9369		BX093694
7 90 8.8 859 13 BQ222616 BQ22 8 90 8.8 987 13 BX357770 BX35 9 8.8 1015 13 BX4046531 BX405 0 90 8.8 1049 29 CNS040AW AL26 1 90 8.8 1201 3 BX39449 BX39 3 90 8.8 1201 9 AL562877 AL56287 BX33 4 90 8.8 1201 13 BX339469 BX33 4 90 8.8 1564 12 BC82 BX33	ž	σ	•	н	13	3492		AGENCOURT
8 90 8.8 987 13 BX357170 BX35 9 90 8.8 1015 13 BX404631 BX40 0 90 8.8 1049 29 CNS040AW AL28 1 90 8.8 1123 13 BX39849 BX39 2 90 8.8 1201 3 AL5642877 AL56262 3 90 8.8 1201 13 BX333469 BX33 4 90 8.8 154 12 BGB2	'n	σ	•	S	13	2261		AGENCOUR
9 90 8.8 1015 13 BX404631 BX40 0 90 8.8 1049 29 CNS040AW AL26 1 90 8.8 1123 13 BX398349 BX39 2 90 8.8 1201 9 AL562877 AL562 3 90 8.8 1201 13 BX339469 BX33 5 90 8.8 1201 13 BX339469 BX33	ñ	On		æ	13	5717		BX357170
0 90 8.8 1049 29 CNSO40AW AL26 1 90 8.8 1123 13 8X398349 BX39 2 90 8.8 1201 9 AL562877 AL562 3 90 8.8 1201 13 BX339469 BX33 4 9 0 8.8 1564 12 BGB27780 BX33 5 9 8.8 1564 12 BGB27780 BX33	35	σ'n		0	13	3463		BX404631
1 90 8.8 1123 13 BX39849 BX39 2 90 8.8 1201 9 AL55429 AL5528 3 90 8.8 1201 13 BX339469 BX33 4 90 8.8 154 12 BG825380 BX33 5 9 6 8 154 12 BG825380 BX33	4	σ	•	9	29	240A		Tetraodor
2 90 8.8 1201 9 AL562877 AL562 3 90 8.8 1201 13 BX339469 BX33 5 80 8.8 1564 12 BG827880 B882 5 80 8 400 14 GB66233	7	6	•	12	13	X39834		BX398349
3 90 8.8 1201 13 BX339469 BX33 5 00 8.8 1564 12 BGB2780 BC82 5 0.5 0 100 11 CBC8730 CBC82	4	6		20	σ	56287	L562	AL562877
4 90 8.8 1564 12 BG827880 BG827880 60275239	4	o	•	20		X33946	BX33	BX339469
5 00 K 0 0 400 14 ABORDS	4	6		56		2788	2788	60275239
	7	, ס ע		2		5030	20.00	A MONTH

ALIGNMENTS

1146 bp mRNA linear EST 12-MAR-2002	~		16182				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;	Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo.	
BM915803		BM915803	BM915803.1 GI:19366182	EST.	Homo saptens (human)	Homo sapiens	Eukaryota; Metazoa;	Mammalla; Eutheria;	1 (bases 1 to 1146)
RESULT 1 BM915803/c LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM			REFERENCE

~

```
147 eSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGl 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3.8
104.50
41.10%
31.29%
10.28%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1. .1031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                       186 oval 187
                                                                                                                                                                                                                                                        681 TTTG 678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                             CB950999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                   g
                                                                                                         ö
                                                                                                                                                     셤
                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                       ò
                                                              Ontpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: OrTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequenced by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
Clone distribution: Agencore story: 256.

High quality sequence story: 256.

High quality sequence story: 256.

Location/Qualifiers

Loc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     859
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90 rThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGln------GlySerAr 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    107 9SerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAs 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      127 pValileProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProil 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       833 -------CGGGAGGCGCGCGGGGCCCGGTAGTGTGGGCCCGGGTCCCCCTCT 787
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61 rSerlleAsnGlyValLeuTrpThrValTyrHis------GlyAlaGlyThr-- 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 AlaGlnGlnThrArgGlyGluGluGlyCysGlnGluThrSer-GlnThrGlyArgAspLy 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         77 -------ArgThrIleAlaSerPro------LysGlyProValThrGlnMetTy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                948 AGCGGAGGGGCGCCGCGTCGGGGCGCCGGCGGGGGGGGCCCG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  902 ------CAGATGTGCGGGTGGGAAGCGCCCGCTCGCGCGGTGGGGGCCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  41 sAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaTh
                            National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1146
47
19
59
57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  858 ACTIGCTIGICGIGITITCIGIGG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-18 (1-197) x BM915803 (1-1146)
     NIH-MGC http://mgc.nci.nih.gov/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3.1
106.00
36.26%
25.82%
10.42%
                                                   Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
```

```
1031 bp mRNA linear EST 29-APR-2003
AEROCOGFT_13445496 NIH_MGC_1177 Mus musculus cDNA clone
CB950999
CB950999.1 G1:30205777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : |||||||:::
786 CGGGTATCTACAGGCGCGGAGGGAGCCACATACCTCTCTCGG------TG 742
                                                      167 yilePheArgAlaAlaValSerThrArgGlyValAlaLysAlaValAspPhe---IlePr 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    395 ATTCAAGGCTATCCCCAAAACAAGAGAGTACATCGGCAGCTTTTCCTT---CACTCATT 451
                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1031)
HH-MGC Http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         44 ValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIle 63
                                                                              þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CDNA Library Preparation: Michael Brownstein Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Apencourt Bioscience Corporation Clone distribution: MGC clone distribution: MGC clone distribution information can bfound through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov column: 11 http://image.llnl.gov b column: 11 High quality sequence stop: 333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism="Mus musculus"
/mol_type="mRNA"
/db_xref="texon:10090"
/clone="IMAGE:30316162"
/lab_host="DH108 (T1-phage-resistant)"
/clone_lib="NIH_MGC_177"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1031
51
16
62
35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-18 (1-197) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus (house mouse)
Mus musculus
```

```
287 CTGACTCTTCTGGGAACCACTGCCCAGCAAAACGGCAGCCCCGGGGCCCATC----
        871 TCAGGGCGTTTTAAAGCCCCCGGCCTTCGCGCCGGGCGAAGCA 913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /sex-"female"
                                                                                                                                                                               AA036834.1 GI:1509872
                                                                                                                                                                                                                     Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         111.6
95.50
38.51
25.00%
9.39%
                                                                                                                                                                                                                                                                                                                                                                                                              Unpublished
Contact: Wilson RK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                              AA036834/c
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                   SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                        TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                             ACCESSION
VERSION
                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                   KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                    /clone_lib="Lupaki_sympathetic_trunk"
/clone_lib="Lupaki_sympathetic_trunk"
/note="Vector: pcMv-SPORT6 (Life Technologies); Site_l:
/note="Vector: pcMv-SPORT6 (Life Technologies); Site_l:
/note_solid: cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
Directionally cloned using the following adaptors:
5'-GACTAGTTCTAGATCGCAGCGCCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
1 kb for average insert length 1.9 kb. This is a primary
1 lbrary, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -----SerProArgProllcSerTyrLeuLys-------GlySerSerGl 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              751 AAAACCCCACCTCGGCCCACCGATGCGCTAAGCCTCCCTTTACAAGCCACGGCGGG 810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       yGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerTh 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           484 TGGGATCCATITITAATAAAGGGTGCTCTGTTAATCATGGCGCCCACGGGCCCGGTGATA 543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           88 GlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArg 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr 87
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://mage-llnl.gov
Plate: LLMA13595 row: c column: 13
High quality sequence start: 57
High quality sequence stop: 394.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     119 LeuTyr-LeuValThr------ArgHisAlaAspValIleProValArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   rArgGlyValAlaLysAlaValAspPheIleProValGluSer 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1199
40
17
54
44
6
                                                                                                                                                                                                                                                                                                          /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    211 t
                                                                                                                                                                                                           /organism-"Homo sapiens"
/mol_type-"mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-18 (1-197) x BQ892487 (1-1199)
                                                                                                                                                                                                                                               /db_xref="taxon:9606"
/clone="lMAGE:6192708"
/sex="male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Technologies."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          11.5
100.50
36.77%
25.81%
9.88%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   175
                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score:
```

ŏ 셤 õ a δ g à a ò a õ 셤 ö 윱 ð

```
EST 26-AUG-1996
                                                                                                                        --- 237
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Not
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2
AA036834 50ares_pregnant_uterus_NDHPU Homo sapiens cDNa clone IMAGE:471945 5' similar to PIR.A55195 A55195 chordin precursor - African clawed frog ;, mRNA sequence.
                                                                                                                                                                                                                                                                                                                     Email: esrewatson.wustl.edu
This clone is available royalty-free through LLNL; contact the
This clone is available royalty-free through LLNL; contact the
TMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -2BM13 rev2 from Amersham
High quality sequence stop: 300.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     34 ThrSerGlnThrGly-----ArgAspLysAsnGlnValGluGlyGluValGlnIleVal
                                                                                                                                                                                                                                                                 Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     502
37
20
45
46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                                                                                                          1. .502
/organism="Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                /mol_type="RNA"
/db_xref="GDB:3757947"
/db_xref="taxon:9606"
/clone="IMAGE:471945"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-18 (1-197) x AA036834 (1-502)
```

S

Oy	52 SerThralaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeu 67		
DP 3	236AGCTGCATGCGTNCNCCAGCTGGGGGTG 207	Alignment Scores Pred. No.:	res:
٥y	68TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerPro 82	Score: Percent Similarity: Doct Local Similarity:	arity:
Dp 3	206 GGCCCCCGACCCCACTGGACACTGTTGCAGCAGTCGGT	Query Match:	יי דטר דווי
	LysGlyProValThrGlnMetTyrT	us-09-965-594-18 (1	-18 (1-
අ	167CAGGCC 138	0v 71	TVIHIS
Oy 1	103 ProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuVal 122	171	::: ::: 171 CACCAC
Db 1	137 AGCCGGGGACACTGCACCTTCTCACAGTGCACCTCTCCAGTGCCCCCCTTGCAGGTGCAG 78		Threlo
Oy 1	123 ThrArgHisAlaAspValIleProValArg-ArgArgGlyAspSerArgGlySerLeuLe 142		
QQ	GCACAA		AST ACAMAG
	142 uSerProArgProlleSerTyr 149	UY 107 Db 285	10/ Argser 285 AGAAGC
	29 GTACCCGTTGCCCGCCACTTC 8	Qy 127	127 Aspval
RESULT 6 BF863244		Db 345	345 GAGATC
LOCUS DEFINITION	BF863244 701 bp mRNA linear EST 19-JAN-2001 963042C02.x1 C. reinhardtii CC-1690, Stress condition I, normalized Tambda 2an II Chiamidomonae unichaeteté nous muss consistentes	QY 139	1
ACCESSION VERSION	, rannon der 11 Chiamiyomohas Lellinaldili Conn, mnn Sequence. BF863244.1 GI:12253388	Db 405	GAAGGG
KEYWORDS	EST.	Ογ 150	150 LeuLys
ORGANISM	Chlamydomouas reinhardtii Chlamydomoas reinhardtii Chlamydomotas reinhardtii	Db 465	465 CTCCAA
	Chlamydomodaceae; Chlamydomonas.	Qy 170	170 Argala
AUTHORS	<pre>1 (Dases 1 to /U1) Grossman,A., Davies,J., Federspiel,N., Harris,E., Hauser,C.,</pre>	Db 525	::: 525 AAAGAA
TITLE	<pre>Letebvre, P., McDermott, J.P., Shrager, J., Sillow, C. and Stern, D. Analyses of the Chlamydomonas reinhardtil Genome: A Model,</pre>	Qy 182	182 ValAsp
	Unicellular System for Analyzing Gene Function and Regulation in Vascular Plants; project phase 3	285 40	GTGGAC
JOURNAL	Unpublished Contact: Charles Hauser	RESULT 7	
	DOMB BOX 91000 Duke University		BF304699
	Durnam, NC 2//08-1000 Tel: 919 613 8159	z	0188825 RNA seq
	rax: 919 013 81// Email: chauser@duke.edu.	z	BF304699 BF304699
FEATURES SOUTCE		KEYWORDS ES	EST. Homo sab
		NISM	Ното зар
	/molype mann /strain="scc1690 wild type mt+ 21gr" /dh vref="tayon:3055"		Eukaryor Mammalia 1 /h=20
	/cloud in the state of the stat	3 S	NIH-MGC
	/note='Vector: pBluescript II SK-; Site_1: BcoR1; Site_2:	AL.	Unpublis
	Anot; into library, constructed by John Davies and Jelfrey McDermott, combines CDNAs from CC-1690 cells grown to mid-log phase in Tab-w (30 min 1hr 4hr) Tabe-s (30 min	COMMENT	Contact: Email: c
	Int. 411, TAP-P (4hr, 12hr, 24hr), NOS to NH4 (30min, 1hr, 4hr) and NH4 to NOS (30min, 1hr, 4hr) Polya mRNA was		CDNA Li
	purified from each sample, pooled and cDNA synthesized. The CDNA was directionally cloned into lambda 2an II		DNA Seq
	(Stratagene) in the EcoRi (5') and XhoRI (3') sites. pBluescript II SK- plasmids were excised from the lambda) [4	found th
	ZAP clones by superinfection with ExAssist (Stratagene) phage. The library was normalized using method 4 described		High qu a
BASE COUNT ORIGIN	173 a 213 c 175 g 140 t	SOLICE	

```
9 984 bp mRNA linear EST 21-NOV-2000
SZF1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4122276 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1-MGC http://mgc.nci.nlh.gov/.
idoal Institutes of Health, Mammalian Gene Collection (MGC)
ublished
ublished
tact: Robert Strausberg, Ph.D.
il: cgapbs-r@mail.nh.gov
iil: cgapbs-r@mail.nh.gov
iil: cgapbs-r@mail.nh.gov
NA Library Preparation: Ling Hong/Rubin Laboratory
NA Library Preparation: Ling Hong/Rubin Laboratory
NA Library Preparation: Ling Hong/Rubin Consortium (LINL)
A Sequencing by: The I.M.A.G.E. Consortium (LINL)
A Sequencing by: The I.M.A.G.E. Consortium (LINL)
A Sequencing by: Loryte Genomics, Inc.
one distribution: MGC clone distribution information can be
in this sequence stop: 646.
Location/Qualifiers

Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                       rLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                llle----- 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         nMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                     CATITGCAIGTGGCTTIAGTCACCCCAAGAAGAGCCTGGGAGTGGGCAITTAIAA 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ------GlySerLeuLeuSerProArgProlleSer---Tyr 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePhe 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :|||||||:::|||
| AGGCAGCAATGGGAGCCITTCGCGGTGTGGGGGTGCTCGTCCTCTAATGTCAGG 524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       aAlaVal------------------SerThrArgGlyValAlaLysAla--- 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ta; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
a; Eutheria; Primates; Catarrhini; Hominidae; Homo.
es 1 to 984)
                                                                                                                                                                                                                 sGlyAlaGlyThrArgThrIleAlaSerProLys-------GlyProVal
                                                                                                                                                                                                                                                                                                                                                    GTTACATACACGG - - - - - AAGGACCAGCGCCTTGGCCACCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                pPhelleProValGluSerLeuGluThrThrMetArg 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | ||| :::||| CGCATACATTGGAAGACACACGGTGCACTCTACGA 626
  701
38
22
52
63
31
  Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1. .984
/organism-"Homo sapiens"
/mol_type-"mRNA"
                                                                                                                                                                     -197) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      equence.
19
19.1 GI:11251586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 piens (human)
17.9
95.50
38.96%
24.68%
9.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          piens
                                                                            ty:
```

Sequencing Center information can be

9

```
1. .629
/organism="Mus musculus"
/mol_type="mRNA"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone='INMAGE:3977578"
/tissue_type="MK cells (flow-sorted)"
/tissue_type="MK cells (flow-sorted)"
/tissue_type="MK cells (flow-sorted)"
/clone_lib="MCI_CGAP_Sp2"
/clone_lib="MCI_CGAP_Sp2"
/note="organ: spleen: Vector: pCMV-SPORT6 (Life
Technologies): mRNA made from flow-sorted NK cells, cDNA
made by oligo-dT priming. Directionally cloned. Average
linsert size 1.5 kb. Primary library, non-amplified. cDNA
Library preparation: David B. Kizaman, Ph.D."
1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91 ThrasnValaspLysaspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeu--- 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        608 GAAGGAGACACACAGATCATCCCTTGTGCA---AAAACATTTCCCATTCCAGATATAAAT 552
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               159 LeucysProAlaGlyHisAlaValGlyIlePheArgAlaAla-------ValSer 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          74 ------AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   500 AAGAGAACTAAAAGCTGTGAGAACCCCTCAAGTTGCCCAGATCCAATGCCCGGGTGAAGCCA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   335 IGCAAAAAACGITICCGAIACTACACTGAGACCACAGAGAITIGTGCTGGAGACTIGAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           45 GluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            440 GNGAATGTG---TGCAGTGTGGCTGGCTGG------GGGTCAAGGTCCATCAAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      110 ---ThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              395 GACACTAAAGCATCTGCCCGCCTGCGAGAGGTTCAACTGGTCATCCAGGAGGACGAGGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     275 ---AAAATAAAGACTCCT-------TTCAAGGGTGACTCGGGGGACCCCTT
  DNA Sequencing by: Washington University Genome Clone distribution: NCI-CGAP clone distribution found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         629
46
22
58
58
46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    175 ThrArgGlyValAlaLysAlaValAspPhellePro 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65 GlyValLeuTrpThrValTyrHisGly-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129 IleProvalArgArgArg------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-18 (1-197) x BG089727 (1-629)
                                                                                                                            Seq primer: -40UP from Gibco
High quality sequence stop: 4
Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17.5
95.00
39.53%
26.74%
9.34%
                                                                                                          MGI:1477610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 9
BU542842/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
           /clone="IMAGE:4122276"
/tissue_type="thabdomyosarcoma"
/tissue_type="thabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="WIH_MGC_17"
/note="Organ: muscle: Vector: poTB7; Site_1: EcoRI;
Site_2: Xhol: CDNA made by oligo-dT priming.
Site_2: Xhol: CDNA made by oligo-dT priming.
Directionally cloned into EcoRY/Xhol Sites selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of Callfornia, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
33 a 329 c 351 g 171 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BG089727 629 bp mRNA linear EST 26-JAN-2001 mab90e06.x1 NCI_CGAP_Sp2 Mus musculus cDNA clone IMAGE:3977578 3' similar to SW:GRAD_MOUSE P11033 GRANZYME D PRECURSOR;, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100 TrpGlnAlaProGlnGlySerArgSerLeuThr---ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         119 LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -------ACCAGGCACGCAACATACATGCAAGAGAGACGTGGT---TCCCGC 554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       139 GlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeu 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         159 LeuCysProAlaGlyHis-Ala-----ValGlyIlePheArg-----AlaAlaVa 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      173 lSerThrArgGlyValAlaLySAlaValAspPheIleProValGluSerLeuGluThrTh 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 629)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             984
39
6
26
7
                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-18 (1-197) x BF304699 (1-984)
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus musculus (house mouse)
Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BG089727.1 GI:12572290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1::|||
427 CAAGGGGGGCGCCC 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              193 rMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                            27.9
95.50
42.86%
37.14%
9.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Tumor Gene Index
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BG089727
                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 8
BG089727/c
                                                                                                                                                                                                                                                                            COUNT
                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMMENT
                                                                                                                                                                                                                                                                            BASE CO
```

a ò 셤 õ 셤

ŏ

a

g

ð g ò

å

158

-----GlyAspSerArg 138

441

^

```
/do.me="INAGE:5742981"
/tissue_type="medula"
/tissue_type="medula"
/lab host="DH108"
/cloom_lib="Will MGC_119"
/cloom_lib="Will MGC_119"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: Not1;
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: Not1;
/note="Organ: brain; Vector size is oligo-dT primed and anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size is destroyed upon cloning). Average insert size is destroyed upon cloning). Research described for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH MGC Library.

43 a 673 c 521 g 288 t 8 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1733 bp mRNA linear EST 20-FEB-2002 ACENCOUR_6558368 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742981 BM553374
                                                                                                                                                                                                                                                                                                                                                                                 584 GCACTGCCGCCCAGCCCTCCATCTCAGCGGGATGTGCAGGGTGAGACAGGAATGCAGGGA 525
                                                                                                                                                                                                       100 TrpGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu 119
                                                                                                                                                                                                                                                                       524 CGTTCTGCCCCTAGGTCAGCCTCTTCATCCGCCTGTTGTGCTTGTCGATGGTCAAGGTTG 465
                                                                                                                                                                                                                                                                                                                                          120 -----TyrLeuValThrArgHisAlaAsp-----ValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              134 Arg------Arg 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 146 ProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGly----- 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     644 CAGTAGCAGCCCCATCCAGAGGAAGACCACTCCGGAGGGCCACAGGCCTCTGCAGGCCTG 585
                                                                   80 AlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGly 99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1733)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Confact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nh.gov
   Tissue Procurement: Life Technologies, Inc.
   CDNA Library Preparation: Life Technologies, Inc.
   CDNA Library Arrayed by: The I.M.A.G.E. Consortium
   CLONE distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LiML at:
   http://image.llnl.gov
   Plate: LiAM12761 row: p column: 22
   High quality sequence start: 88
   High quality sequence start: 88
   High quality sequence start: 88
   High quality sequence stop: 539.
   Location/Loulifiers
   Location/Lules**
   //mol.type***
   //mol.type***
   //mol.type***
   //mol.type***
   //mol.type***
   //mol.type***
   //mol.type***
   //mol.type***
   //mol.type***
   //mol.type**
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 164 -----HisAlaValGly 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     284 CGGGTGCTCACGGTGCAGGG 264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BM553374.1 GI:18792049
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 243 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
BM553374/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                qq
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       l. 905
//organism="Homo saplens"
//organism="Homo saplens"
//mol_type="mana"
//db_xref="taxon:9606"
//dbc="TyAGE:6574789"
//tissue_type="carcinoma, cell line"
//lab_host="DH108 (phage-resistant)"
//lab_host="DH108 (phage-resistant)"
//clone_lib="NNH_MGC_40"
//orde="Corgan: prostate: Vector: poTB7; Site_l: XhOI;
//ore="Corgan: prostate: Vector: poTB7; Site_l: XhOI;
//site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhOI sites using the following 5' adaptor: GGCAGGGGG(3. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."
BU542842 905 bp mRNA linear EST 13-SEP-2002 AGENCOURF_10334715 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6574789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 04 AAGTGGGAGGCAGGCCATGGTGCACCTGGGGAGGCCCCTGGTGCAGAAGCAGCCCCA 645
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28 GluGluGlyCysGlnGluThrSerGln --- ThrGlyArgAspLysAsnGlnValGluGly 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------GlyThrArgThrIle 79
                                                                                                                                                                                                                                                                   Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 905) NIH-MGC http://mgc.nci.nih.gov/. NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2770 row: k column: 13
High quality sequence stop: 633.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     47 GluValGlnIleValSerThrAlaThrGlnThrPheLeuAla----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61 -----ThrSerIleAsnGlyValLeuTrpThr-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-18 (1-197) x BU542842 (1-905)
                                                                   5', mRNA sequence.
BU542842
BU542842.1 GI:22853325
                                                                                                                                                                                                          Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       39.5
93.50
31.88%
24.64%
9.19%
                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                              Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                            ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
                            DEFINITION
                                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
```

ò 요 ŏ g à

ò g œ

õ a ò đ à a ö g ò g ò g ò g à g ò 8 õ

```
/dev_stage="adult"
/dev_stage="adult"
/clone_lib="An expressed sequence tag (EST) collection
from the resurrection plant Selaginella lepidophylla"
/note="Vector: Lambda Uni-Zap XR, Bluescript SK-; Site_1:
EOSTI: Site_2: Xhoi: Library construction was performed
according to manufacture's (Stratagene, Inc.) recommended
protocol for the Lambda UniZapXR vector and cDNA synthesis
                                                                                                                                                                                                                                                                                                                                                                                /tissue_type="microphyll fronds undergoing desiccation for 2.5 h"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AlahlaValSerThrArgGlyValAlaLysAlaValAspPhe---IleProValGluSer 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123 ------ThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ||| ::: |||||||||||| :::
230 ---TCAGCCGCTAAIGGICGICCAAICCAAGACGIGAICCAGACAGAIGCCGCIAITAAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             287 CCTGGGAACAGCGGGGGTCCGCTATTGGACAGTTCTGGAAATTTGATAGGCATCAACAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113 CTIGGAAGIICGICCGAICIGCIIGIIGGCCAGAAGGIGIAIGCIAICGGIAAICCIIII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         173 GGATTGGATCATACGCTGACCACAGGCGTCATCAGTGGTCTTCGAAGGGAGATTACT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139 GlySerLeuLeuSerProArgProlleSerTyrLeu--------------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               528
115
126
26
26

    .528
    /organism-"Selaginella lepidophylla"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CysGlySerSerAspLeuTyrLeuVal------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      137 t
                                                  Department of Blochemistry
University of Newada
MS200, Reno, NV 8957-0014, USA
Tel: 775-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
PCR PRIMERS
FORWARD: T3 20mer
BACKWARD: T7 21mer
Plate: 005 row: F column: 12
Seq primer: T3 20mer
                                                                                                                                                                                                                                                                                                                                    /mol_type="mRNA"
/db_xref="taxon:59777"
/clone="SLA005F12"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-18 (1-197) x BM402566 (1-528)
                                                                                                                                                                                                                                                              High quality sequence stop: 528 Location/Qualifiers
Selaginella lepidophylla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      137 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21.9
93.00
42.98%
30.58%
9.14%
                                    Contact: Cushman JC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      125 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                  Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GTT 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..
⊙
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                  JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12
                                                                                                                                                                                                                                                                                    FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ôγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               BM402566 528 bp mRNA linear EST 01-JUL-2002 SLA005F12_34513 An expressed sequence tag (EST) collection from the resurrection plant Selaginella lepidophylla Selaginella lepidophylla cDNA clone SLA005F12 5, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAla 171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGATGCTGCAGGAGTGCCCCGGGCTGCCCAGTGGCTGACGGTGGAAAGTGCCCACGTGCT 548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AlavalSer-----ThrArgGlyvalAlaLys------AlaValAspPheIle 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Selaginella lepidophylla
Selaginella lepidophylla
Eukaryota; Viridiplante; Streptophyta; Embryophyta; Tracheophyta;
Lycopodiophyta; Isoetopsida; Selaginellales; Selaginellaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100 -----------TrpGlnAlaProGlnGlySerArgSerLeuThrProCys 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           133 ArgArgGlyAspSerArgGlySerLeuLeuSerProArg---ProIleSerTyrLeuLys 151
                                                                                                                                                                                                                                                                851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80 AlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGly 99
                                                                                                                                                                                                                                                                                                                                                                                62
                                                                                                                                                                                                                                                                                                    [ (bases 1 to 528)
Ituriaga, G. and Cushman, J.C.
An expressed sequence tag (EST) collection from the resurrection
                                                                                                                                                                                                                                                                                                                                                                              45 GluGlyGluValGln----IleValSerThrAlaThrGlnThrPheLeuAlaThrSer
                                                                                                                                                                                                                                                                                                                                                                                                                    907 CGGGGGGACCAAATICGCIGCITGTICCCAATIGCACCAAGGCI---GIITCAIICCAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CGCCGGGGGTGATGGCCGCCCAGGGCTGGCCCATCCTTCCCTGGCCCTGGAGAGGGCCAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  790 CCCTCCCGGGAGGGTGAGCTCCAAAGGGGGTTTTTCC--------CTCCCAGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   113 ThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                          63 IleAsnGlyValLeu----TrpThrValTyrHisGlyAlaGlyThrArgThr---Ile
                                                                                                                                                                                                                              -----GlyGluGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             745 recegrecegecerretreggaagececaggecaggaaga----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         186 ProvalGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                              GlyAspThrAlaTyrAlaGlnGlnThrArg-----
                                                          Length:
Matches:
                                                                                                                                                                                        (1-1733)
                                                                                                                                                                                        US-09-965-594-18 (1-197) x BM553374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BM402566.1 GI:21643782
                                                        92.2
93.50
37.26%
26.42%
9.19%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Selaginella.
                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         607
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             487
                                                                                                                                                                                                                              11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
BM402566
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE
AUTHORS
TITLE
```

셤

σ

```
/organism="Sorghum bicolor"
/wol_type="genomic DNA"
/wol_type="genomic DNA"
/wol_type="genomic DNA"
/done='ic83bil"
/lab_host="JM107 or DH5a"
/clone='ic83bil"
/clone='ib='w65-SbicolorF (JM107 adapted methyl filtered)"
/clone='ib='w65-SbicolorF (JM107 adapted methyl filtered)"
/clone='ib='w65-SbicolorF (JM107 adapted methyl filtered)"
/note="Site_l: Xba I; Site_2: Xba I; The vector was
digested with Xba In one nucleotide was added by fill in
in the recessive 3' end. The genomic DNA was nebulized,
end repaired, adaptor ligated and size fractionated using
sephadex. The resulting fragments were between 0.8 and 3
kb and were cloned into the vector (x/y reads in M13mpl9,
-b/g reads in pUC19: The same ligation was transformed in
either JM107 or DH5a."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----CysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   119 LeuTyrLeuvalThrArgHisAlaAspValIleProValArgArgArgGlyAsp----- 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    320 CACCCAGTGCCCCCGGCGGCCGCGTACGGGGACAGGAAACTGGTGTGGTGAGGATCAC 261
                                                                                                                                                                                                    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PacCAD clade; Panicoideae; Andropogoneae; Sorghum.

1 (bases 1 to 701)
Rabinowicz,P.O., O'Shaughnessy,A.L., Balija,V., Dedhia,N., Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L., Zutavern,T., Palmer,L., McCombie,W.R. and Martienssen,R.A. Genomic shotgun sequences from Sorghum bicolor (methyl-filtered) Dopublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        440 GIGGCCGIGGCCCGCCGCCCCCCCCCCGCGTGGAATTAIGGCCGCAATAAACTCGGCAC 381
1c83b11.b1 WGS-SbicolorF (JM107 adapted methyl filtered) Sorghum bicolor genomic clone ic83b11 5', genomic survey sequence. BZ342381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 AlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LeuValGlyTrpGlnAlaProGlnGlySerArg --- SerLeuThrProCysThr ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 -----IleAlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Contact: W. Richard McCombie
Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Tel: 516 367 8884
Exa: 516 367 8874
Email: mccombieGeshl.org
Plate: ic83 row: b column: 11
Seq primer: -21M13UnivPwd
Class: shotgun
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      701
48
14
55
51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      High quality sequence stop: 701
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-18 (1-197) x BZ342381 (1-701)
                                                                                                                                                             Sorghum bicolor (sorghum)
Sorghum bicolor
                                                                                               BZ342381.1 GI:24742983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31.7
93.00
36.69%
28.40%
9.14%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1. .701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
                                                                                                                                                             SOURCE
ORGANISM
      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                       REFERENCE
AUTHORS
                                                                    ACCESSION
                                                                                               VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           임
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                       Submitted (27-Ann-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Unpublished
This sequence was generated from the T7 end of BAC 283L13. 283L13 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GSS 06-NOV-2002
                                    GSS 29-JAN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  345 AGCAGCTGGTCATGTCGGTGGTCCAGCAGTCTTGGAAAGGAC-------TGGAGATCC 295
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 111 | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgArgGlyAspSerArg 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                234 TIATITICIGATIACCATGGGGAGGAGTIGTCGCACGAGGAGGCCTTAGGGAGACGGGAGG 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  139 GlySerLeuLeuSerProArgPro-----IleSerTyrLeuLysGlySerSerGlyGly 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           157 ProLeuLeu---CysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThr 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ---ProGlnGlySerArg-----SerLeuThrProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TGGACAGCATCCTGGCAGAGTGGGAGCTCTCGAGGGCGCTGTCCGCCGTCACCAGTGACC 346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlnMetTyr-----ThrAsnValAspLysAspLeuValGlyTrpGlnAla 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28
                                                                                                                                                                                                                                               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostel; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

I (bases 1 to 644)

Humphray,S.J., Huckle,E. and Durham,J.L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         117 CCCTTTGCCTGGATGCCCAGTCCTGGCCCTGCCCTGTCACCCGGGCATGGAAACCTTCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThr
             өлдэвчв 644 bp DNA linear GSS 29-JAN-
Danio rerio genomic clone DKEY-283L13, genomic survey sequence.
BX238988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  644
37
118
52
52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            l. 644
/organism="Danio rerio"
/organism="Tonomic DNA"
/mol type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-283113"
/tissue_type="Testis"
/note="vector pIndigoBAC-536"
]? c 176 g 127 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  http://www.sanger.ac.uk/Projects/D_rerio/
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   701 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-18 (1-197) x BX238988 (1-644)
                                                                                                                               BX238988.1 GI:28161322
                                                                                                                                                                                              Danio rerio (zebrafish)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28.4
93.00
44.72%
30.08%
9.14%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        176 ArgGlyval 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAAGGCTTA 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                           Danio rerio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BZ342381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  B2342381/c
                                LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNT
                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
TITLE
JOURNAL
      8X238988/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 13
                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE CO
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SOCOS
```

a

셤 ò

q ò

à g ò g ò g ð g

ò

```
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
 US-09-965-594-18 (1-197) x BU541777 (1-1213)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
1. .772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CC406704.1 GI:30886794
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Other_GSSs: PUHKD12TD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Class: sheared ends
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seq primer: TR
                                                                                                                                                                                                                              114 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          180 LysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11 CGCTCT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2ea mays
2ea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CC406704/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                           염
                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                               a
                                                                                                                            qq
                                                                                                                                                                õ
                                                                                                                                                                                             셤
                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/vualifiers

1. .1213
/organism="Homo saplens"
/mol_type="mrnh"
/db_xref="taxon:9606"
/clone="IMAGE:6573516"
/tissue_type="carcinoma, cell line"
/lab_host="DHIOB (phage-resistant)"
/clone_lib="NIH_MGC_40"
/clone_lib="NIH_MGC_40"
/clone_lib="nih made by oligo-dT priming.
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRi/xhoI sites using the following 5' adaptor: GGCAGAG(6). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library.
75 a 341 c 463 g 134 t rary.
                                                                                                                                                                                                                                                                                                        AGENCOUR_10325396 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6573516 BU541777
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bloscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://mage.llni.gov
Plate: LLCM2767 row: f column: 12
High quality sequence stop: 165.
                                                  215 GTTTGACGCGCGCGCGCCGCGGCCGCGGCACCTCTGCTCTGCTCGGCCGTGGGAAAAA 156
                                                                                              150 LeuLysGly-----LeuCy 160
                                                                                                                                                                160 sproAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyVal---- 178
GIGGAGAGGCCAGGACGACGACGA-------CGACGACGACGGGGGGACCCCG 216
                                  -----SerArgGlySerLeuLeuSerProArgProIleSerTyr 149
                                                                                                                               Eukaryota; Metazoa: Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria: Primates; Catarrhini; Hominidae; Homo. (Dases 1 to 1213)
NIH-MGC http://mgc.nci.nih.gov/.
National institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1213
41
13
50
50
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                              179 -AlaLysAlaValAspPheIlePro 186
                                                                                                                                                                                                                                                             47 rececrecerrecrerecre 23
                                                                                                                                                                                                                                                                                                                                                                                           BU541777.1 GI:22852258
                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64.9
93.00
38.038
28.878
9.148
                                                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
260
                                137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                              RESULT 14
BU541777/c
                                                                                                                                                                                                                                                                                                                               LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORIGIN
                                                                                                                       g
                                                                                                                                                                                                                                                           셤
                                õ
                                                               셤
                                                                                              ò
                                                                                                                                                                ò
                                                                                                                                                                                             셤
                                                                                                                                                                                                                           õ
```

```
CC406704 772 bp DNA linear GSS 19-MAY-2003 WHKD12TB ZM_0.6_1.0_KB Zea mays genomic clone ZMABTa469B24, genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /db_xref="taxon:e>/,
/clone="ZMMBTa469B24"
/clone_11b="ZM_0.6_1.0_KB"
/rote="Vector: pCR+TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT =="elected genomic DNA library"
CoT == 208 t
162 AlaGlyHisAlaval-----GlyIlePheArgAlaAlaValSerThrArgGlyValAla 179
                                                                                                                                                                                    -----CysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                   127 AspValileProValArgArgArg-------GlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                      142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                              GlnMetTyrThrAsnValAspLysAspLeuValGly-----TrpGlnAlaPro--- 103
                                                                                                                      353 TTGTTGTATTTGTCTCGCAGGTCATTCCAAGTCGGCTGCATCCGCTGGCAGTGTCCACAC 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            122 CGCCGCCGCCATGGCCCAGCAGCAGCAGCAGCAGCAGTCAGGGCCG------CCG 72
                                                                                                                                                             104 GlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843
Fax: 301-838-0208
```

ORIGIN					
Alignment Scores: Pred. No.: Score: Percent Similarity: Dest Local Similarity: Query Match:	res: arity: milarity:	40.3 92.50 40.578 92.108 29.108	Length: Matches: Conservative: Mismatches: Indels: Gaps:	772 47 24 59 46 10	
US-09-965-594	-18 (1-197	US-09-965-594-18 (1-197) x CC406704 (1-772)	-772)		
Qy 39 Db 633	Argasplys.	AsnGlnvalGlud ::: AACGACGTAAACAAG	GlyGluValGlnIleV GGCTATGTTACAATTG	39 ArgAspLysAsnGlnValGluGlyGluValGlnIleValScrThrAlaThrGlnThr ::: ::: 633 AGAGGGGGAACGRAGGGCTATGTACAATTGTGACGAGCCCTATG	57 580
Oy 58	PheLeuAla::: TACGAGGCC	ThrSerlleAsnGly :::::: 	ValLeuTrpThrV ::: TATTCTATCTGGAGG-	PheLeualaThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThr :::	76 538
Oy 77 Db 537		AlaSerProLysGly) ::: CTGACGATGAGATCA	ProValThrGlnMet1 AGGGTGACACAGATGA	ArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAsp 96 ::: :: ::: :::	96 484
Oy 97 Db 4 83	LeuValGly' AGGACCCAG	TrpGln2 -	AlaProGlnGlySerA CAATAACTGGGATTAC	LeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 	113
Oy 114 Db 423		SerAspLeu TCCTCAACGGTCAGG	Tyri AGCTGCTGGCCCTACT	CysGlySerSerAspLeu	127
Qy 128 Db 363	ValilePro ACTTCATCA	ValArgArgArgGly; ACTGCAAGGCGGCA-	AspserArgGlyserI CAACAGCGTGGGGTT	128 ValileProValargArgGlyAspSerArgGlySerLeuLeuSerProArgProlle 147	305
Oy 148 Db 304	SerTyrLeu	LysGlySerSerGly ::::: AGGTGGCAAATGGA	GlyProLeuLeuCysF ::: GACCCAGTTTTCTGCC	148 SerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGly :::::: ::::: :::::	167
Qy 168 Db 259	::: GTAACTCGT	AlaAlaValSerThr GCGCA	IlePheArgAlaAlaValSerThrArgGlyValAlaLysAlaValAspPhe ::: GTAACTCGTGGGGAGCCATTGACATCAA	hrargglyValAlaLysAlaValAspPhe	184
Qy 185 Db 223	AAGTTCACC	ATTGAGGCATATGCA	185	ceuglu 191 TTGAG 179	

Search completed: August 31, 2003, 04:27:41 Job time: 1917.31 secs

```
(BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M, Weinheimer
  86.5
882.5
  10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PX X X X X X Y X Y
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                             August 30, 2003, 17:42:58 ; Search time 44.6227 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
5: /SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
6: /SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*
7: /SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*
8: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*
8: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*
8: /SIDSI/gcgdata/geneseqy-embl/AA1980.DAT:*
8: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1980.DAT:*
8: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA1990.DAT:*
8: /SIDSI/gcgdata/geneseqy/geneseqp-embl/AA2000.DAT:*
                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Description
                                                                                                                                                                                                                                                                                                                                                                                                                                               1107863
                  GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                          Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                              1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                          OM protein · protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15224
AAB15225
AAB15223
AAB15222
AAB15222
AAB15220
AAB15220
AAB15220
AAB15220
AAB15220
AAB15220
AAB15220
                                                                                                                                                                                                                                                                                                                               BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A_Geneseq_19Jun03
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        length: 0
length: 2000000000
                                                                                                                                                                                                                                   US-09-965-594-20
1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              197
197
197
197
197
195
195
665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            100.0
99.0
98.5
97.1
92.7
92.7
92.1
89.4
86.8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB seq
DB seq
                                                                                                                                                                                                                                            Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1020
1010
1005
990
973
946
939
912
885.5
                                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                   Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Database
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Minimum Maximum
                                                                                                                                                    Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4500000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Š
```

02 500 65.3 66.3 50	NS4A-NS3
878 5 86.1 500 2.0 878 5 86 1 516 50	
878.5 R6.1 665 20	
877.5 86.0 665 20	HCV NS4A-NS3 COMPI
875.5 85.8 216 20	
874.5 85.7 216 20	
874.5 85.7 665 20	
873 5 85 6 671 20	HCV NS4A-NS3 COMPI
871.5 85.4 216 20	HOV NOAB-NOS COMPL
870.5 85.3 216 20	
870.5 85.3 665 20	HCV NS4A-NS3 compl
870.5 85.3 671 20	
870 85.3 215 20	
867.5 85.0 216 20	
05 017 020 65/00	HCV NS4A-NS3 COMPI
02 012 0.00 0.000	
863.5 84.7 216.20	HOW NOAB-NOS COMPL
863.5 84.7 216.20	
859 84.2 213 20	
859 84.2 631 20	q
858.5 84.2 191 21	Henatitis C wing
858.5 84.2 3011 19	Hebatitis C virus
858.5 84.2 3011 24	າ ຕ
858.5 84.2 3012 23	Cvirus
855.5 83.9 3011 14	
854.5 83.8 687 16	coded
854.5 83.8 1648 16	
854.5 83.8 1766 10	e encoded
854.5 83.8 1786 10	sedneuce
854.5 83.8 2261 10	encoded
854.5 83.8 2301 10	ednence encoded
44 634.5 83.8 2436 10 AAPYZUSU 45 854.5 83 8 2436 10 AAPYZUSU	Sequence encoded 1
	בוורחתבת
ALIGNMENT	STN
RESULT 1	
815224	
ID AAB15224 standard; protein; 197 AA.	
AC AAB15224;	
10.00-2000	
25111	
DE Hepatitis C virus NS4A-NS3 fusion proteas	ease *6.
KW Hepatitis; NS3 protease; viral replication KW liver failure; liver cancer; mutant; mutei	tion; chronic liver disease;
OS Hepatitis C virus. OS Synthetic.	
PN WO200040707-A1.	
AA PD 13-JUL-2000.	
XX PF 06-JAN-2000; 2000WO-US00345.	
XX PR 08-JAN-1999; 990S-0115271.	
ALCC.	

Modified hepatitis C virus (HCV) NS3 protease comprising at least 1

Goldfarb V;

Zhang Y,

s,

WPI; 2000-465976/40. N-PSDB; AAA73333.

```
ô
                                                                                                                                                                                             The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the MEV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural guidies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helixo-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 TSINGVLMTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            121 LVTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MKKKGSVVIVGRINLSGDTAYAQOTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA 60
substitution of a hydrophobic alpha-helix \theta amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Repatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100.0%; Score 1020; DB 21; Length 197; 100.0%; Pred. No. 3.2e-98; ative 0; Mismatches 0; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease #7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ż
                                                                                                                                  Claim 23; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB15225 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 100. Matches 197; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
N-PSDB; AAA73334.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB15225;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15225
AAB
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
ö
                                                                                                         The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and proteins.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61 TSINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA 60
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening Inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                      Query Match 99.0%; Score 1010; DB 21; Length 197; Best Local Similarity 99.5%; Pred. No. 3.6e-97; Matches 196; Conservative 0; Mismatches 1; Indels 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinhelmer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ş
                                                                                 Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                         197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAA73332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15223;
                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

0 The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the mutaviral treatments of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:Inhibitor complexes. This sequence contains the alpha-helix0-1 120 121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180 9 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHIGRDKNQVEGEVQIVSTATQTFLA Gaps Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein. ö Length 197; Indels Weinheimer S, Zhang Y, Goldfarb V; 98.5%; Score 1005; DB 21; 98.5%; Pred. No. 1.2e-96; Hepatitis C virus NS4A-NS3 fusion protease #4 2; Mismatches Z (BRIM) BRISTOL-MYERS SQUIBB CO. Claim 23; Fig 15; 66pp; English. AAB15222 standard; protein; 197 AVDFIPVESLETTMRSP 197 AVDFIPVESLETTMRSP 197 06-JAN-2000; 2000WO-US00345 99US-0115271 (first entry) Conservative Query Match Best Local Similarity Matches 194; Conserv 197 AA; Hepatitis C virus WO200040707-A1. Wittekind M, 08-JAN-1999; 19-DEC-2000 13-JUL-2000 Synthetic Sequence 181 181 AAB15222; RESULT 4 õ a à g g q ò

WPI; 2000-465976/40.

```
0
                                                                                                                                      The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the try genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3 NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                            substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                               protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                             61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                               protease comprising at least 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                              ..
0
                                                                                                                                                                                                                                                                                                                                                                               Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                              4; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                               Score 990; DB 21;
Pred. No. 4.4e-95;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease #3.
                                                                                                                                                                                                                                                                                                                                                                                                           2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Zhang Y,
                            Modified hepatitis C virus (HCV) NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                          Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAB15221 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                             97.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      990S-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                             Matches 191; Conservative
                                                                                                                                                                                                                                                                                                                                                                                              Similarity
                                                                                                                                                                                                                                                                                                                                               197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
N-PSDB; AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15221;
                                                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                   variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           XXEEEEXXXXCCCCCCCCCCCXXXXCCCCCCCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                đ
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                 ö
                                                                                                        The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                       TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                                                                                                                                                                                                        TSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                          9
                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                  LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK
                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3 protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                           Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Goldfarb V;
                                                                                                                                                                                                                                                                         95.4%; Score 973; DB 21;
95.4%; Pred. No. 2.6e-93;
11ve 2; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  liver failure; liver cancer; mutant; mutein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhang Y,
                                                                                        Claim 23; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAB15226 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                      AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                      Local Similarity 95.4 es 188; Conservative
 2000-465976/40.
                                                                                                                                                                                                                                                     197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
           N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic
                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAB15226;
                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                    Ω
                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
ö
                                                                                                                                                                                                                                                                               The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HVV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix {\bf 0} amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             92.7%; Score 946; DB 21; 94.4%; Pred. No. 1.7e-90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus NS4A-NS3 fusion protease #2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15220 standard; protein; 195 AA
                                                                                                                                                                                                                                   Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99us-0115271.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 94.4
16; Conservative
2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
                            N-PSDB; AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W0200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ma.
Local St.
186;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
```

us-09-965-594-20.rag

```
N-PSDB; AAA73328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-JUN-1999
                                                                                                                                                                                                                                                        complexes.
                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY24943;
                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY 24943
  ò
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                        61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                             119 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAK 178
                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                         Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                            1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                         ä
                                                                                                                                                                                                                                                                                                  Length 195
                                                                                                                                                                                                                                                                                                                       Indels
 Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                               92.1%; Score 939; DB 21;
93.4%; Pred. No. 9.2e-90;
live 3; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                            Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15212 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDEIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AVDFIPVESLETTMRSP 197
 s,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                             Best Local Similarity 93.4 Matches 184; Conservative
Wittekind M, Weinheimer
                     WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                             195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                             Claim 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15212;
                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
 ð
                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 ISINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antivixal treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-HS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:Inhibitor
                                                                                                                                                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDKDLVGWPAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
      Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 912; DB 21;
Pred. No. 6.1e-87;
1; Mismatches 12;
   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY24943 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                                                     Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETTMRSP 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 92.4%;
Matches 182; Conservative
Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      07-SEP-1999 (first entry)
                                                                 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
Synthetic.
```

9

```
VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US2452B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                              Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (SCHE ) SCHERING CORP.
(SCHE ) SCHERING CORP.
                                                             WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-JUL-1998;
28-NOV-1997;
                            Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Synthetic.
                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24942;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY24942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    201
                                                                                                                                                                  The present invention describes a covalent hepatitis C virus (HCV) MS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 earlie protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTAIQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 GSVVIVGRINLSGD -- - TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                             э;
т
                                                                                                                                                                                                                                                                                                                                                                                                                DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                             Match 86.8%; Score 885.5; DB 20; Local Similarity 85.7%; Pred. No. 2e-83; es 168; Conservative 15; Mismatches 10;
                                             Yao N;
                                                                                                         New hepatitis C virus covalent complexes
                                                                                                                                         Claim 6; Page 90-92; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY24947 standard; Protein; 665 AA
                                             Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV NS4A-NS3 complex SEQ ID NO:18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                          WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24947;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                202
                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Loca
Matches
```

셤

ŏ g

g å qq ò

```
62 SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        82 CVNGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .22 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present invention describes a covalent hepatitis C virus (HCV) NS4ANS1 complex comprising a central hydrophobic domain of native HCV NS4A peptide, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4ANS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4ANS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        S GSVVIVGRINLSGD --- TAYAQQTRGEQGCQRTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 86.5%; score 882.5; DB 20; Length Best Local Similarity 85.2%; Pred. No. 4.1e-83; Matches 167; Conservative 16; Mismatches 10; Indels
Yao N;
                                                                                                                                                                                    New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                           Claim 6; Page 100-102; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAY24942 standard; Protein; 665 AA.
Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV NS4A-NS3 complex SEQ ID NO:13.
```

```
Weber PC, Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24946 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       970S-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Malcolm BA, Taremi SS,
 Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP.
                                 WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                               Sequence 216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-JUN-1999
 Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24946;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY 24946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                            The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex. and an HCV NS3 Serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5 GSVVIVGRINLSGD - - - TAYAQQIRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62 SINGVLWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL
                                                                                                                                                                                                                                                                                                                                                                                                                                  3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                 DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                 86.4%; Score 881.5; DB 2985.7%; Pred. No. 5.2e-83; tive 14; Mismatches 11
                   Yao N;
                                                                                 New hepatitis C virus covalent complexes
                                                                                                                  Claim 6; Page 88-90; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY17880 standard; Protein; 216 AA.
                 Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                 Local Similarity 85.7%
Les 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                 Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                  665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-1999
                 Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
```

Matches

õ 셤 ò a ò ద a

```
62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                               NS4A.NS3 complex comprising a central hydrophobic domain of native NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5 GSVVIVGRINLSGD---TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                   The present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
86.1%; Score 878.5; DB 20; Length 216;
Best Local Similarity 85.6%; Pred. No. 2.2e-83;
Matches 167; Conservative 15; Mismatches 10; Indels 3;
YAO N;
                                                                                                                                                                       New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                          Claim 6; Page 76-77; 211pp; English.
```

Sun Aug 31 15:59:42 2003

```
WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY17884
       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 VTRHADVIPVRRRGDSRGSLLSPRPISYI,KGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 GSVVIVGRINLSGD---TAYAOQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            / Match 86.1%; Score 878.5; DB 20; Length 665; Local Similarity 85.2%; Pred. No. 1.1e-82; nes 167; Conservative 15; Mismatches 11; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Weber PC, Yao N;
                                                                       New hepatitis C virus covalent complexes
                                                                                                                   Page 97-99; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV NS4A-NS3 complex SEQ ID NO:12.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY24941 standard; Protein; 665
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-0524528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (SCHE ) SCHERING CORP.
                      WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY 24941;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                      Claim 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
AAY24941
ID AAY24941
XX
XX
XX
XX
DDT 07-SE
DT 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 셤
  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
```

```
62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                    The present invention describes a covalent hepatitis C virus (HCV) NS4A NS3 complex comprising a central hydrophobic domain of native HCV NS4A Peptide.

NS4A peptide, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tebbered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ArPass activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5 GSVVIVGRINLSGD -- - TAYAQQIRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQIFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCRITSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 86.0%; Score 877.5; DB 20; Length Best Local Similarity 85.2%; Pred. No. 1.4e-82; Matches 167; Conservative 15; Mismatches 11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yao
New hepatitis C virus covalent complexes
                                                           Claim 6; Page 85-87; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY17884 standard; Protein; 216 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV NS4A-NS3 complex SEQ ID NO:8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    182 VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAY17884;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4 aperide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                         New hepatitis C virus covalent complexes
                                                                               Claim 6; Page 80-81; 211pp; English.
```

216 AA; Sequence

1; 62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121 122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181 5 GSVVIVGRINLSGD---TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT 61 81 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT 3; Gaps Query Match 85.8%; Score 875.5; DB 20; Length 216; Best Local Similarity 85.1%; Pred. No. 4.6e-83; Matches 166; Conservative 16; Mismatches 10; Indels 3; g g õ q ò õ ò

Search completed: August 30, 2003, 19:12:24 Job time : 44.6227 secs

g

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

using sw model protein search, OM protein August 30, 2003, 19:02:22; Search time 16.2134 Seconds (without alignments) 1168.492 Million cell updates/sec ü o

US-09-965-594-20 Perfect score:

1 MKKKGSVVIVGRINLSGDIA.....VAKAVDFIPVESLETIMRSP 197

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

Sequence:

283308 seqs, 96168682 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database

PIR_76:*
1: pir1:*
2: pir2:*
1: pir3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Query Result

Description	qenome polyprotein							genome polyprotein	genome polyprotein	genome polyprotein	genome polyprotein	polyprotein dour	polyprotein - marm	probable aromatic	probable periplasm	hypothetical prote	nitrate/nitrite se	nitrate/nitrite se	hypothetical prote	Q	low density lipopr	probable oxidoredu	htrA-like serine p	beta transducin-li	probable sugar upt	hypothetical prote	~	hypothetical prote	glycosyl transfera
DI CI	GNWVC3	S40770	GNWVCH	GNWVTW	A45573	GNWVIC	GNWVCJ	S18030	JC5620	GNWVJ8	JQ1303	T08841	T08839	H83144	B71284	G83612	B81104	C81911	A75335	139383	JE0372	H70847	н97199	T42045	F95973	B71360	A49448	T04533	D87603
8	_	m	-	~	-	7	_	Н	Ţ	ч	٦	N	ď	~	~	N	7	~	~	(1	7	7	~	~	N	N	~	~	~
Match Length	3011	3011	3011	3010	3010	3010	3010	3010	3014	3033	3033	3005	2970	209	398	716	290	290	377	452	1615	479	433	1049	394	495	764	846	322
Match	83.8	83.7	83.2	82.0	81.1	80.7	80	79	72	99	99		7						8.1	8.1	8.1	7.9	7.8	7.8	7.7	7.7	7.7	7.7	7.6
Score	854.5	853.5	848.5	836.5		823.5	823.5	811.5	743.5	675	673	251	245	85.5	85.5	84.5	83.5	83.5	83	83	83	80.5	79.5	79.5	78.5	78.5	78.5	78.5	78
No.	1	7	m	∢	S	φ	7	æ	σ	10	Π	12	13	7	15	16	17	18	19	20	21	22	23	24	25	56	27	5 8	53

ä

..

17; Indels

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST 53

ð

Query Match 83.8%; Score 854.5; DB 1; Length 3011; Best Local Similarity 82.8%; Pred. No. 9.7e-70; Matches 169; Conservative 9; Mismatches 17; Indels 9;

hypothetical prote	HIV-1 retropensin	hypothetical prote	proteinase (EC 3.4	mosA protein - Rhi	probable exported	hypothetical prote	transferrin-bindin	qenome polyprotein	qenome polyprotein	genome polyprotein	hypothetical prote	hypothetical prote	hypothetical prote	requiatory protein	hypothetical prote
E75392	GNLJGL	S57895	PRLJHD	B53308	AH1030	A81942	F81196	GNWVY	GNWVYP	GNWVNE	T32434	G90654	G85505	E64745	B83477
~	-	N	~	-	~	~	~	٦	~	_	~	7	~	~	7
477	1165	361	323	333	492	539	915	3411	3411	3414	424	452	452	452	515
7.6	7.6	9.7	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.4	7.4	7.4	7.4	7.4
78	78	77.5	76.5	76.5	76.5	76.5	92	76	92	76	75.5	75.5	75.5	75.5	75.5

```
genome polyprotein - hepatitis C virus (strain HCV-1)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein Ns4s; nonstructural protein NS4s; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 19-Jan-2001
C;Coate: 30-Sep-1992 #text
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Actatus: preliminary
A; Status: preliminary
A; Molecule type: genomic RNA
A; Restus: preliminary
A; Molecule type: genomic RNA
A; Restus: preliminary
A; Molecule type: genomic RNA
A; Resperimental source: isolates E-b17
C; Superfamily: hepatitis C virus genome polyprotein
F; 1-115/Product: capsid protein C *status predicted <NED>
F; 116-191/Product: nonstructural protein Bi status predicted <NED>
F; 730-1006/Product: nonstructural protein NS1 *status predicted <NS1>
F; 730-1006/Product: nonstructural protein NS2 *status predicted <NS2>
F; 1230-1237/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif B
F; 1312-1317/Region: protein motif Bidding motif B
F; 1316-1319/Region: protein motif Bidding motif B
F; 1316-1319/Region: protein motif Bidding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: mRNA
A; Molecule type: mRNA
A; Molecule type: mRNA
A; Residues: 1-3011 < CHO>
A; Cross-references: GB: MG2321; NID: 9329873; PIDN: AAA45676.1; PID: 9329874
A; Cross-references: GB: MG321; NID: 9329873; PIDN: AAA45676.1; PID: 9329874
B; Chan. Virol. 73, 1131-1141, 1992
A; Tile: Analysis of a new hepatitis C virus type and its phylogenetic relationship the A; Accession: PQ0403
A; Accession: PQ0403
A; Accession: PQ0403
A; Molecule type: genomic RNA
A; Residues: 1577-1633 < CHA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: DDBJ:D10128
A;Experimental source: isolates E-b16
A;Accession: PQ0404
```

~

```
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                             genome polyprotein - hepatitis C virus (strain H) N; Contains: capsid protein C; envelope protein H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A. Cross references: GR: D00831: NID: 9221511; PIDN: BAA00705.1; PID: 9221512

A. Experimental source: isolate HC-J1

A. Experimental source: isolate HC-J1

C. Superfamily: hepatitiss C virus genome polyprotein

C. Superfamily: hepatitiss C virus genome polyprotein

C. Superfamily: hepatitiss C virus genome polyprotein

C. Superfamily: hepatitiss C virus predicted corporation

F. 2-115-192. Product: apsid protein C * status predicted corporation

F. 116-191. Product: major envelope protein B * status predicted corporation

F. 192-389/Product: monstructural protein NS1 * status predicted corporation

F. 1900-1615/Product: nonstructural protein NS2 * status predicted corporation

F. 1900-1615/Product: nonstructural protein NS2 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NS2 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG2 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG2 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG2 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG2 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG2 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation

F. 1910-1615/Product: nonstructural protein NG3 * status predicted corporation corporation corporation corporation corporation corporation corporation corporat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
N;Contains: capsid protein C; envelope protein MS4s; nonstructural protein NS5s; nonstructural protein NS4s; nonstructural protein NS5scies: hepatitis C virus
C;Species: hepatitis C virus
C;Accession: $40770; PC1285
R;Okamoto, H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A;Cross-references: EMEL:D10749; NID:g221586; PIDN:BRA01582.1; PID:g221587
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60; 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
A;Reference number: PC1284; MUID:91013116; PMID:2170712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1005 RKGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
.005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                   54 ATOTFLATSINGVLWTVYHGAGTRTIASPKGPVTOMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 853.5; DB LE NO. 1.2e-69;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1) Score 853.5; allarity 82.8%; Pred. No. 1.2e Conservative 8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           submitted to the EMBL Data Library, March 1992 A; Reference number: $40770 A; Accession: $40770 A; Accession: $40770 A; Accession: $40770 A; Accelle type: genomic RNA A; Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                     1185 CTRGVAKAVDFIPVESLETIMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA
A; Residues: 1-513 <OK2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local Similarity
Matches 169; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Accession: PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                              à
                                                                                                                                                                                                                                                                                                                 aa
                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                   ď
                                                                                                                                                    q
```

```
A; Molecule type: genomic RNA
A; Residues: 1-3011 < INC>
A; Residues: 1-3011 < INC>
A; Cross-references: GB: M67463; NID: 9329737; PIDN: AAA45534.1; PID: 9329738
A; Cross-references: GB: M67463; NID: 9329737; PIDN: AAA45534.1; PID: 9329738
A; Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, W.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus: comp A; Reference number: A41546; MUID: 92052256; PMID: 1658800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A:Note: neither amino acid nor nucleotide sequence is given
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein
C:Superfamily: hepatitis C virus genome polyprotein; divorbitis protein; glycoprotein; hydrolase; nonstruct
C:Stepards: ATP; capsid protein; envelope protein c *status predicted <CPC>
F:115/Product: capsid protein C *status predicted <CPC>
F:130-139/Product: major envelope protein M *status predicted <MSI>
F:390-729/Product: nonstructural protein NSI *status predicted <NSI>
F:100-1615/Product: nonstructural protein NSI *status predicted <NSI>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             genome polyprotein - hepatitis C virus (strain Taiwan)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 4sequence_revision 31-Dec-1992 ttext_change 19-Jan-2001
C;Accession: A36814; A41546
R;Inofauge, G:Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
a;Dentited to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               F:1516-1862/Product: nonstructural protein NS4m *status predicted <N4A>
F:1863-2013-Product: nonstructural protein NS4b *status predicted <NN4B>
F:2014-3011/Product: nonstructural protein NS4 *status predicted <NN5P
F:2014-3011/Product: nonstructural protein NS5 *status predicted <NN5P
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1065 ATOTFLATCINGVCWTVYHGAGTRTIASPKGPVIQTYTNVDQDLVGWPAPQGSRSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C:Species: heptitis C virus
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A4024
R:Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 848.5; DB 1; Length 3011;
Pred. No. 3.5e-69;
10; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  83.2%;
81.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity ....
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F,1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                            A;Reference number: A36814
A;Accession: A36814
```

m

```
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 76.5
Matches 156; Conservative
                                                                                                                                                     Query Match
Best Local Similarity
Matches 157; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Species: hepatitis C virus
C; Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C; Accession: A45573, 1992
A; Title: Wolecular cloning of hepatitis C virus genome from a single Japanese carrier: 8
A; Reference number: A45573, MuiD:92295714; PMID:1318627
A; Accession: A45573
A; Reference cumber: A45573, MuiD:92295714; PMID:1318627
A; Accession: A45573
A; Reference cumber: A45573; MuiD:92295714; PMID:1318627
A; Accession: A45573
A; Reference cumber: A45573
A; Reference: BR:D11168; GB:D01171; NID:9221612; PIDN:BAA01943.1; PID:9221613
A; Restdues: 1-3010 CTAN>
A; Ross-references: GB:D11168; GB:D01171; NID:9221612; PIDN:BAA01943.1; PID:9221613
A; Ross-references: GB:D1168; GB:D0171; May are dicted < Ross-
F; 191-389/Product: monstructural protein NS1 *status predicted < NS2>
F; 130-1237/Region: nucleotide-binding motif A (P-loop)
F; 131-137/Region: nucleotide-binding motif B (F-loop)
                                                                                                                                                            A; Cross-references; GB:M84754
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein; glycoprotein; hydrolase; nonstructura
C; Superfamily: hepatitis C virus genome protein; glycoprotein; hydrolase; nonstructura
E;1-115/Product: capsid protein C *status predicted <CPC>
F;116-191/Product: envelope protein M *status predicted <ARE>
F;192-189/Product: major envelope protein B1 *status predicted <ARE>
F;390-729/Product: nonstructural protein NS1 *status predicted <NS1>
F;130-1105/Product: nonstructural protein NS1 *status predicted <NS2>
F;1307-115/Product: nonstructural protein NS2 *status predicted <NS2>
F;1307-137/Region: nucleotide-binding motif A (P-loop)
F;1317-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (EC 3.4.21.98) (nonstructu
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping A:Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNOVEGEVOVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1065 ATQSFLATCINGVCWTVYHGAGSKTLAGPKGPITQMYTNVDQDLVGWHAPQGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      genome polyprotein - hepatitis C virus (strain JT)
N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78.4%; Pred. No. 4.4e-68; tive 18; Mismatches 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82.0%; Score 836.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 78.4
Matches 160; Conservative
                                                                                                                   A:Molecule type: genomic RNA A:Residues: 1-3010 <CHE>
                                                                                     A; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstrancentein NS44); nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            C;Accession: A38465
R;Takamizawa, A.; Mozi, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, J. Yirol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from hu A;Reference number: A38465; MUID:91140698; PMID:1847440
A;Reference number: A38465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      nonstruc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A Readule type: genomic RNA
A Readules: 1-3010 caRA
A Residues: 1-3010 caRA
C Superfamily: hepatitis C virus genome polyprotein
C Superfamily: hepatitis C virus genome polyprotein
C Superfamily: hepatitis C virus genome polyprotein
C Superfamily: hydrolase; nonstructing
F; 2-115/Product: capsid protein C *status predicted <REN>
F; 192-389/Product: major envelope protein B *status predicted <NES>
F; 390-729/Product: nonstructural protein NS1 *status predicted <NS2>
F; 300-1006/Product: nonstructural protein NS2 *status predicted <NS2>
F; 130-1237/Region: nucleotide-binding motif A (P-loop)
F; 1312-1317/Region: nucleotide-binding motif B
F; 1316-1319/Region: nucleotide-binding motif B
F; 1316-1319/Region: nucleotide-binding motif B
F; 1316-1319/Region: DRSH motif
F; 1603-2013/Product: nonstructural protein NS4a *status predicted <NAB>
F; 1863-2013/Product: nonstructural protein NS4b *status predicted <NAB>
F; 1863-2013/Productin NS4b *status predicted <NAB>
F; 1863-2013/Productin NS4B **
F; 1864-1865/Productin NS4B **
F; 1865-1865/Productin NS4B **
F; 1865-1865/Productin NS4B **
F; 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN------LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1005 RRGKEILLGPADSLEGRGLRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1065 ATQSFLATCVNGVCWTVYHGAGSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                   Gaps
F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                            81.1%; Score 827.5; DB 1; 77.0%; Pred. No. 2.9e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   80.7%; Score 823.5; DB 1;
76.5%; Pred. No. 6.8e-67;
tive 21; Mismatches 18;
                                                                                                                                                                                                                                                                                                                  20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRGVAKAVDFIPVESLETTMRSP 197
```

ă

```
genome polyprotein - hepatitis C virus (isolate EUR1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstruprotein NS4s; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;bate: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Richamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Blochem. Biophys. Res. Commun. 236, 44-49, 1997
A; Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomins
A; Reference number: JC5620; MUID:97366593; PMID:9223423
                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
A; Residues: 1-3010 <HON>
A; Residues: 1-3010 <HON>
A; Crossreferences: EMBL:X61596; NID:q59478; PIDN:CAA43793.1; PID:q59479
A; Crossreferences: EmbL:X61596; NID:q59478; PIDN:CAA43793.1; PID:q59479
A; FHonda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163.169, 1993
A; Title: Sequence analysis of putative structural regions of hepatitis C virus isolat A; Reference number: A48332; MUID:93119270; PMID:8380322
          C; Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 23-Mar-2001 C; Accession: S18030; S33570; A48332; S18029 R; Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S. submitted to the EMBL Data Library, September 1991 A; Description: A whole genome of hepatitis C virus cDNA was isolated from a single A; Reference number: S18028 A; Accession: S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1125 YGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAV 1184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               79.6%; Score 811.5; DB 1;
ilarity 76.0%; Pred. No. 8.7e-66;
Conservative 20; Mismatches 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local Similarity
nes 155; Conserv
      A; Variety: isolate JK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A Molecule type: genomic RNA
A; Residues: 2650-2707 <kA2>
C; Comment: The cleavage sites of this polyprotein have not been determined.
C; Comment: The cleavage sites of this polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
F; 2-115/Product: envelope protein M *status predicted <RDE>
F; 1922-899/Product: monstructural protein NS1 *status predicted <NS2>
F; 1300-726/Product: nonstructural protein NS2 *status predicted <NS2>
F; 1300-155/Product: nonstructural protein NS2 *status predicted <NS2>
F; 1316-1319/Region: nucleotide-binding motif A (P-loop)
F; 1316-1319/Region: DEXH motif
F; 1861-2013/Product: nonstructural protein NS4 *status predicted <N4S>
F; 2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F; 196, 209, 234, 250, 305, 325, 417, 423, 430, 448, 532, 556, 576, 623, 645, 1213, 1255, 2041, 2077, 2240, 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome polyprotein - hepatitis C virus (isolate JK1)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: _hepatitis C virus
                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructura
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.Molecule type: genomic RNA
A.Residuaes: 1-3010 < kARA:
A.Straduaes: 1-3010 < kARA:
A.Straduaes: 1-3010 < kARA:
A.Strato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A.Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence variance on the control of the non-A, non-B hepatitis viral genome show sequence variance on the control of the non-A, non-B hepatitis viral genome show sequence variance on the control of the non-A, non-B hepatitis viral genome show sequence variance on the control of the non-A, non-B hepatitis viral genome show sequence variance on the control of the control of the non-A, non-B hepatitis viral genome show sequence variance number:
                                                                                                                                                                                                                                                                                                                                                                                                                            R;Kato, N.; Hljikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimot Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients A;Reference number: A39253; MUID:91088550; PMID:2175903
A;Accession: A39253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1005 RRGKEILLGPADSFGEGGWRLLAPITAYSQQTRGLLGC11TSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 ATOTFLAISINGVLWTVYHGAGIRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRINLSGD-----TAYAQOTRGEOGCOKTSHTGRDKNOVEGEVOIVST 53
                                                                                                                                                                                                                                                                                                                               C;Species: hepatitis C virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 19-Jan-2001
C;Accession: A39253; PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80.7%; Score 823.5; DB 1; Length 3010; 75.5%; Pred. No. 6.8e-67; Live 23; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETTMRSP 197
174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 75.5;
Matches 154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: PS0086
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
   ò
                                                                 g
```

S

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19 TAYAQOTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match 66.2%; Score 675; DB 1, Best Local Similarity 69.8%; Pred, No. 3e-53; Matches 125; Conservative 24; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ouery Match 66.0%
Best Local Similarity 68.7%
Matches 123; Conservative
                                                                  <KAT>
                         A; Molecule type: mRNA A; Residues: 2678-2729
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
C. Species: hepatitis C virus
C. Species: hepatitis C virus
C. Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C. Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C. Accession: A40250: p00397: p00559
R. Accession: A40250
R. A. Marcante R. C. A. Marante R. C. Marante R. Marante R. Marante R. C. Marante R. Mar
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A:Experimental source: isolate E-bl2
R:Kato, N: Octsuyama, Y:: Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimotohno
Biochem Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720309
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (strain HC-J8)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4a; nonstructural protein NS4a; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A: Molecule type: genomic RNA
A: Residues: 1.3033 < CNA>
A: Residues: 1.3033 < CNA>
K: Residues: 1.3033 < CNA>
K: Cross-references: GB: D10988; GB: D01221; NID: g221608; PIDN: BAA01761.1; PID: g221609
R: Chau, S.W.: McOmish, P.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, P.I.
J. Gen. Virol. 73, 1131-1141, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to A:Reference number: PQ0393; MUID:92268871; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <u>,..</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1006 RRGREIFLGPADDIKTSGWRLLAPITAYAQQTRGVLGAIVLSLTGRDKNEAEGEVQFLST 1065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114 CGSSDLYLYTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 72.9%; Score 743.5; DB 1; Length 3014; Best Local Similarity 68.6%; Pred. No. 1.5e-59; Matches 140; Conservative 25; Mismatches 30; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Cross-references: DDBJ:D10134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A: Molecule type: genomic RNA A: Residues: 2678-2754 <CHA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Accession: PQ0397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

Gaps

ö

```
A.Molecule type: genomic RNA
A.Residues: 1-3033 < CORA-
A.Residues: 1-3034 < CORA-
A.Residues: 1-3034 < CORA-
A.Residues: 1-3034 < CORA-
A.Residues: 1-3034 < CORA-
A.Experimental source: 150.1ate HC-J6 from a Japanese individual
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein; serine proteinase; tra
F.2-115/Product: capsid protein C *status predicted < CPC>
F.116-131/Product: major envelope protein M *status predicted < CPC>
F.130-389/Product: major envelope protein E *status predicted < CPC>
F.390-733/Product: nonstructural protein NSI *status predicted < NSI>
F.34-1010/Product: nonstructural protein NSI *status predicted < NSI>
F.3116-1321/Region: nucleotide-binding motif B
F.3316-1321/Region: DEXH motif
genome polyprotein - hepatitis C virus (isolate HC-J6)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural strotein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural brotein NS4s; nonstructural protein NS4s; nonstructural protein NS5
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Accession: JQ1303
R; Okamoto, H; Okada, S; Sugiyama, Y; Kurai, K; Iizuka, H; Machida, A.; Miyakawa, J; Gen. Virol. 72, 2697-2704, 1991
A; Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a A; Accession: JQ1303
A; Accession: JQ1303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1620-1866/Product: nonstructural protein NS4a *status predicted <N4A>
F;1867-2017/Product: nonstructural protein NS4b *status predicted <N4B>
F;2018-3033/Product: nonstructural protein NS5 *status predicted <N05>
F;2018-3033/Product: nonstructural protein NS5 *status predicted <N05>
F;196,209,234,305,325,417,423,430,448,477,534,542,558,578,649,1091,1217,1259,2038
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   66.0%; Score 673; DB 1;
68.7%; Pred. No. 4.6e-53;
tive 27; Mismatches 29;
```

OY 53 TATQTFLATSINGVLWTVYHGAGTRTIASPR	RESULT 14 H83144 Probable aromatic acid decarboxylase PA4019 [C;Species: Pseudomonas aeruginosa C;Species: Pseudomonas aeruginosa C;Sacession: H83144 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoa adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.; Lory, S.; Olson, M.V. Nature 406, 959-964, 2000 A;Title: Complete genome sequence of Pseudomon, R; Reference number: A82950; MUID::20437337; PM A;Status: preliminary A;Molecule type: DNA A;Stratus: preliminary A;Molecule type: DNA A;Cross-reference GR.A;F004818; GR.A;F00401.	A; Experimental source: strain PAO1 C; Genetics: A; Gene Family: dedF protein C; Superfamily: dedF protein Cuery Match B 44; Score 85.5; Watches 51: Conservative 16; Mismatch	43 QVEGEVQ-IVSTATQTFLA 1 1 1 1 1 1 1 1 1 1	Oy 138 RGSLLSPR*-PISYLKGSSGGPLLC	RESULT 15 971284 probable periplasmic serine proteinase DO (ht probable periplasmic serine proteinase DO (ht c;Species: Treponema pallidum subsp. pallidum C;Date: 24-Jul-1998 #sequence_revision 24-Jul c;Accession: B71284 R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M. rson, J.; Khalak, H.; Richardson, D.; Howell,	Cience 281, 375-388, 1998 A:Title: Complete genome sequence of Treponen A:Reference number: A71250; MOID:98332770; PA:Accession: B71284 A:Status: preliminary; nucleic acid sequence A:Molecule type: DNA A:Residues: 1-398 <ccl></ccl>
19 TAYAQQTRGEQGCOKTSHTGRDKNQVEGEVQIVSTATOTFLATSINGVLWTVYHGAGTRT 78	RESULT 12 108841 108841 108841 108841 108841 108841 108841 108841 10991 10919protein - douroucouli hepatitis GB virus A 1019protein - douroucouli hepatitis GB virus A 1019protein - douroucouli hepatitis GB virus A 10194 10194 10196 10197 10197 10197 10198 10	Query Match 24.6%; Score 251; DB 2; Length 3005; Best Local Similarity 34.1%; Pred. No. 2.6e-14; Matches 56; Conservative 29; Mismatches 69; Indels 10; Gaps 3; Qy 33 KTSHTGRDKNQVEGEVQIVSTAIQTFLATSINGVLWTVYHGAGTRIASPKGPVTQMYTN 92	QY 93 VDKDLVGWQAPQGSRSLTPCTGGSSDLYLVTRHADVIPVRRGDSRGSLLSPRPISYLKG 152 [:	RESULT 13 T08839 polyprotein - marmoset hepatitis GB virus A C;Species: 13-301-1999 #sequence_revision 23-301-1999 #text_change 17-Nov-2000 C;Accession: T08839 R;Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I. J; Gen. Virol. 79, 41-45, 1998 A;Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.	A; Reference number: 216486; MUID:98120818; PMID:9460920 A; Accession: T0839 A; Status: translated from GB/EMBL/DDBJ A; Molecule type: genomic RNA A; Residues: 1.2970 CERK> A; Cross references: EMBL:AF023424; NID:92828597; PIDN:AAC40501.1; PID:92828598 C; Superfamily: hepatitis C virus genome polyprotein C; Keywords: polyprotein	Query Match 24.0%; Score 245; DB 2; Length 2970; Best Local Similarity 27.8%; Pred. No. 9.2e-14; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Qy 3 KKGSVVIVGRINLSGDTAYAQQTRCEQGCQKTSHTGRDKNQVEGVQIVS 52 :: :: :: :: :: ::

```
M.; White, O.; Sutton, G.G.; Dodson, R.; G
1, J.K.; Chidambaram, M.; Utterback, T.; M
J.C.
                                                                                                                                                                                                                                                                                                                                                                                                  monas aeruginosa PA01, an opportunistic pe
PMID:10984043
                                                                                                                                                                                                                                                                                                                                cogucht, S.D.; Warrener, P.; Hickey, M.J.;
S.N.; Folger, K.R.; Kas, A.; Larbig, K.; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; NID:99950200; PIDN:AAG07406.1; GSPDB:GN
                                                                                                                                                                                                                                                               [imported] - Pseudomonas aeruginosa (stra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ema pallidum, the syphilis spirochete. PMID:9665876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5; DB 2; Length 209;
1.7;
ches 61; Indels 55; Gaps 11;
LCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S-----PRPISYLKGSSGGPLICP 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POAMQAFLTEYCGAAAGQI-----RVFG 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----WIVYHGAGTRILASPKGPVTOMYT 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 e not shown; translation not shown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   htrA-1) - syphilis spirochete
um (syphilis spirochete)
ul-1998 #text_change 09-Dec-2002
                                                                                                                                                                                                                                                                                             sp-2000 #text_change 31-Dec-2000
                                                                                                                                                              :| |:: | |
LDREAITHTEAPP 1154
                                                                                                                                      FIPVESLETTMRSP 197
```

```
A,Cross-references: GB:AE001248; GB:AE000520; NID:g3323074; PIDN:AAC65740.1; PID:g332308
A;Experimental source: strain Nichols
C;Genetico:
C;Genetico:
A;Gene: TPO773
C;Superfamily: Escherichia coli trypsin-like proteinase degS; GLGF domain homology; tryp
                                                                                                                69; Gaps 10;
                                                                                                                                                        59 LATSINGVLW-----TVYHGAGTRTIASPKGPV------TQMY------90
                                                                                  Query Match

8.4%; Score 85.5; DB 2; Length 398;
Best Local Similarity 23.0%; Pred. No. 3.7;
Matches 44; Conservative 29; Mismatches 49; Indels 6
                                                                                                                                                                                                                                                                                                           180 KAVDF-IPVES 189
                                                                                                                                                                                                                                                                                                                            | | :||::
255 SGVGFAVPVDT 265
                                                                                                                                                               a 50 a
                                                                                                                                                                                                                                               oy
oy
                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                            ò
```

Search completed: August 30, 2003, 19:20:31 Job time : 17.2134 secs

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

(without alignments) 949.524 Million cell updates/sec August 30, 2003, 18:01:52; Search time 9.75674 Seconds Run on:

US-09-965-594-20 1020 Title: Perfect score:

1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197 BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table: Sequence:

127863 segs, 47026705 residues Searched: Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		æ			SUMMARIES	
Result		Query				
Ž.	Score	Match	Length	80	QI	Description
,	54.	٠.	3011	-	POLG_HCV1	у чер
7	848.5	83.2	3011	-	POLG_HCVH	h genome
m	36.		3010	-	POLG_HCVTW	h genome
4	27.		3010	Н	POLG_HCVJT	h genome
Ś	23.		3010	~	POLG_HCVBK	h genome
9	23.		3010	-	POLG_HCVJA	h genome
7	\sim		3033	~	POLG_HCVJ8	h genome
œ	\sim		3033	~	POLG_HCVJ6	h genome
σ,	m		321	٦	HOA_ARATH	arabidops
10	85.5		209	Н	PAAD_PSEAE	
11	ω,		437	Н	DEG1_ARATH	
12	83		452	-	AAMP_HUMAN	
13	78.5		485	-	Y136_TREPA	
14	ω.		764	-	ICCR_DROME	-
15	78		1165	-	POL_GALV	_
16	Ľ.	•	263	~	GRAK_MOUSE	
17	76.5		323	-	VPRT_SMRVH	
18	ė	•	333	Н	MOSA_RHIME	
19	97	٠	401	~	FXH1_MOUSE	O88621 mus musculu
20	94	7.5	3411	П	POLG_YEFV1	P03314 y genome po
21	97		3411	-	POLG_YEFV2	
22	_	•	3414	-	POLG_TBEVW	t genome
23	75.5		248	-	TRY1_CHICK	
24	'n	7.4	452	-	MLTD_ECOLI	
25	Š.		2269	-	WDR9_HUMAN	Q9nsi6 homo sapien
56	_		467	-	NX1B_BOVIN	
27	•		248	-	GRAD_MOUSE	P11033 mus musculu
28	÷		3414	-	POLG_LANVT	P29837 1 genome po
53	74		911	-	TB11_NEIMB	neisseria
30	74	•	973	-	VP18_HUMAN	09p253 homo sapien
31	7		3414	٦	POLG_TBEVH	t genom
32	73.5	7.2	248	н.	TRY2_CHICK	ıs ga
33	m.		264	Н	CTRL_HUMAN	P40313 homo sapien

HCV_capsid.

MEROPS; 229.001; -.
MEROPS; U39.001; -.
InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C

EMBL; M62321; AAA45676.1; -. PIR; A39166; GNWVC3.

PIR; A39166; GNWVC3. PDB; IAIV; 16-FEB-99. PDB; IHEI; 25-NOV-98.

	P26596 lactococcus	067531 aquifex aeo	P27425 equus cabal	P59240 mus musculu	P28863 oryctolagus	Q28146 bos taurus	P07720 t genome po	004538 t genome po	P29326 hepatitis e	P33426 hepatitis e	POB169 bos taurus
DPM1_USTMA	MCP_BPF41	FLII_AQUAE	TRFE_HORSE	NPH4_MOUSE	MM03_RABIT	NX1A_BOVIN	POLG_TBEVS	POLG_POWVL	VST2_HEVBU	VST2_HEVPA	MPRI_BOVIN
н	~	~	-	7	~	-		-	_	-	-
294	301	443	106	1425	478	1530	3412	3415	9	099	2499
7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2	۲.۲	7.1	7.1
۲.	3.5	73.5	73.5	73.5	73	73	73	73	72.5	72.5	72.5
73	7										

ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PESCENCE FROM N.A.

ESQUENCE FROM N.A.

ESQUENCE FROM N.A.

MEDLINE-91172826; PubMed-1848704;

MEDLINE-91172826; PubMed-1848704;

Choo O.-L., Richman K.H., Han J.H., Berger K., Lee C., Dong C.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J., Weiner A.J.,

Callegos C., Coit D., Medina-Selby A., Barr P.J.,

Callegos C., Coit D., Medina-Selby C., Coit D., Callegos C., Coit D., Callegos C., Coit D., 
                                                                                                                           01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
15-5E-2003 (Rel. 24, Last annoctation update)
66-00me polyprotein (Contains: Capaid protein C (Core protein) (P22);
67-009 (GR09) (GR01); Protein F1 (GP32) (GP35); Envelope glycoprotein E2 (GP63) (GP35); Envelope glycoprotein E2 (GP63) (GP35); Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/hellcase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
NS5E (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate i) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -!- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN H AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND KRNA.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                          PRT; 3011 AA
                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepacivirus.
NCBI_TaxID=11104;
                                                                          POLG_HCV1
RESULT 1
POLG_HCV1
```

```
PRT; 3011 AA.
                                                                                                                                                                                                                                        174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             of unwinding.";
Structure 6:89-100(1998)
                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Repacivirus
                                                                                                                                                                                                                                                                                                                                                     POLG_HCVH
P27958;
                                                                                                                                                                                                                                                                                                                   RESULT 2
POLG_HCVH
                                                                                                                                                                             g
                                                          ò
                                                                                                 a
                                                                                                                                       ò
                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE.
CAPESID PROTEIN C (POTEWITAL).
MAJOR ENVELOPE PROTEIN E (POTEWITAL).
NAJOR ENVELOPE PROTEIN E (POTEWITAL).
NONSTRUCTURAL PROTEIN NS1/EZ (POTEWITAL).
PROTEASE/HELICASE NS3 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS2 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWITAL).
NONSTRUCTURAL PROTEIN NS4B (POTEWITAL).
NONSTRUCTURAL PROTEIN NS4B (POTEWITAL).
NONSTRUCTURAL PROTEIN NS4B (POTEWITAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC. .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 854.5; DB ]
Pred. No. 2.5e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC,
N-LINKED GECONAC,
N-LINKED GEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BOX
                                                                        InterPro; IRR004109; HCV_NS3.
InterPro; IPR001045; HCV_NS4.
InterPro; IPR001046; HCV_NS4.
InterPro; IPR001046; HCV_NS4.
InterPro; IPR001046; HCV_NS5a.
InterPro; IPR001056; HCV_RARP.
InterPro; IPR001050; Helicase_C.
InterPro; IPR001095; RNA_POL_DS_ES.
InterPro; IPR007094; RNA_POL_DS_ES.
InterPro; IPR007094; RNA_POL_DS_ES.
InterPro; IPR007094; RNA_POL_DS_ES.
InterPro; IPR007094; RNA_POL_PSVIr.
Pfam; PF01543; HCV_COTE; I.
Pfam; PF01543; HCV_COTE; I.
Pfam; PF01559; HCV_NS1; I.
Pfam; PF01560; HCV_NS2; I.
Pfam; PF01006; HCV_NS3; I.
Pfam; PF01006; HCV_NS4a: I.
Pfam; PF001001; HCV_NS4a: I.
Pfam; PF001001; HCV_NS4a: I.
Pfam; PF00271; helicase_C; I.
Pfam; PF000271; helicase_C; I.
PRODOM; PE016602; HCV_NS1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DECH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645 645 N-LIN
2041 2041 N-LIN
2240 2240 N-LIN
2364 2364 N-LIN
2362 2363 N-LIN
3011 AA: 327197 MW; G
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                tch 83.8%; al Similarity 82.8%; 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1165
1237
1319
                                          IPR002531;
                                                          IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
Matches 169
                    InterPro;
                                          InterPro;
                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
```

3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST 53

ò

```
1065 AAQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVDQDLVGWPAPQGSRSLTPCT 1124
113
                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                       01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last amonation update)
6600me polyprotein (Contains: Capsid protein C (Core protein E C (GP68) (GP70) (NS1); Protein E1 (GP32) (GP55); Envelope glycoprotein E2 (GP69) (GP70) (NS1); Protein P7; Monstructural protein NS2 (P21) (EC 3.4.99.-); Protease/helicase NS3 (P70) (HeppolyIrin)
(EC 3.4.99.-); Protease/helicase NS3 (P70) (HeppolyIrin)
(EC 3.4.99.); Nonstructural protein NS5A (P65); Nonstructural protein NS5B (P66) (P70) (NNA-directed RNA Polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -I FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-I- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4B-NS5A, AND NS5A-NS5B, NS4A-NS4B, PAG-NS4B, AND NS5A-NS5B, PUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACTIVATION OF NG3.
-!- FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-!- FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN ESSENSTIAL ROLE IN THE VINOS REPLICATION.
-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
Kim - 1.L., Morgenstern K.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1'.

CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ပ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657
```

m

```
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 848.5; DB 1;
Pred. No. 9.2e-72;
                               (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                        (GLCNAC.
(GLCNAC.
                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83.2%;
81.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1417
1420
1436
1439
1453
1463
1468
1488
1488
11501
1501
1511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1385
1393
1409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     353
361
366
368
375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11312
113212
113213
113313
113313
113313
113313
113313
113413
114413
114413
114413
11480
11480
11480
11480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TURN
STRAND
TURN
HELIX
TURN
TURN
TURN
TURN
STRAND
TURN
HELIX
TURN
STRAND
HELIX
TURN
HELIX
HELIX
TURN
HELIX
TURN
HELIX
                               CARBOHYD
TURN
HELIX
TURN
HELIX
TURN
HELIX
TURN
STRAND
TURN
STRAND
TURN
TURN
STRAND
TURN
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
        CARBOHYD
        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation her Burgeran Bioinformatics institute. There are no restrictions on its use by non-profit institutions as 10ng as 1ts content is in no wodified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                 EMBL; M67463; AAA455341; -.

DR PRB; A36844; GNWCH.

DR PDB; HAIR: 73-007-98.

DR PDB; LAIX: 17-UNH-98.

REROPS; S29 001; -.

DR TRANSFCT 704152; -.

DR InterPro; IPR001522; HVV_capsid.

DR InterPro; IPR002519; HVV_carv.

DR InterPro; IPR002519; HVV_carv.

DR InterPro; IPR002519; HVV_carv.

DR InterPro; IPR0014109; HVV_MS1.

DR InterPro; IPR0014109; HVV_MS3.

DR InterPro; IPR001409; HVV_MS3.

DR InterPro; IPR001450; HVV_MS3.

DR InterPro; IPR001550; Helicase_C.

DR InterPro; IPR001569; HVV_MS3.

DR InterPro; IPR001765; HVV_MS4.

DR InterPro; IPR001709; RNA_pol_DS_PS.

DR InterPro; IPR001709; RNA_pol_DS_PS.

DR Ffam; PF01539; HVV_MS1; 1.

DR Ffam; PF01539; HVV_MS1; 1.

DR Ffam; PF01001; HVV_MS4; 1.

DR Ffam; PF01005; HVV_MS4; 1.

DR Ffam; PF01001; HVW_MS4; 1.

DR Ffam; PF01001; HVW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
ATP (POTENTIAL).
DECH BOX.
N LINKED (GLCNAC. ..) (POTENTIAL).
N LINKED (GLCNAC. ..) (POTENTIAL).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY 039. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REMOVED FROM CAPSID PROTEIN C BY CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
NONSTRUCTURAL PROTEIN NS5B.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                191
383
746
809
1026
11057
11711
11711
1107
1108
1108
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
11165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
1165
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         192
384
747
810
1027
1058
1172
2421
347
1083
11165
11165
1230
1316
234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CCHAIN
CC
```

```
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INITUMET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IRANSMEM
   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation he Burpan Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as its content is no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFHAAV 173
                                                                                                            54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                  01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 42, Last annotation update)
Genome polyprotein (Contains notation update)
Genome polyprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4); Nonstructural protein
NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (1solate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KKGSVVIVGRIN-----LSGDTAYAQOTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (RNA)(N).
SUBDNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSIGTS OF TWO PROTEINS:
PROTEIN A AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
 Indels
 18;
10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                             174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWVTW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410; DEAD
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PDB; 1N64; 25-FEB-03.
PDB; 1NS3; 08-APR-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; $29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus.
NCBL_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                    POLG_HCVTW
P29846;
                                                                                                                                                                                                                                                                                                                                                  RESULT 3
POLG_HCVTW
TO 1-APR N
POLG_HCVTW
TO 11-APR N
DT 11-SEP P
DT 11-SEP P
DT 11-SEP P
DE GENOONG
DE GENOONG
DE GENOONG
DE GENOONG
DE NASAB G
OC VITUSE
CC VITUSE
CC 1- CNS
C
                                                                                                                                          a
                                                                                                                                                                                ò
                                                                                                                                                                                                                                                        à
```

```
ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROFEIN C BY THE CELLULAR AMINOEPTIDASE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROYEASE/RELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSA (POTENTIAL).
NONSTRUCTURAL PROTEIN NSAB (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            82.0%; Score 836.5; DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED (GLCNAC...)
N-LINKED (GLCNAC...)
N-LINKED (GLCNAC...)
N-LINKED (GLCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2788 N-LINKED (GLCNAC. . . AA; 327047 MW; AAD267D55CDFE215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1.3e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. 1.3e
18; Mismatches
                                              LOST.

LOST HROUNSS.

LICO, IPROD145; HCV_NS4a.

LICEPPO, IPROD169; HCV_NS4b.

LICEPPO; IPROD2166; HCV_NS5a.

INTERPO; IPRO02166; HCV_RGRP.

DR INTERPO: IPRO07095; RN_pol_DS_PS.

INTERPO: IPRO07094; RNA_pol_DS_PS.

Pfam; PF01543; HCV_caps.id; 1.

Pfam; PF01542; HCV_core; 1.

Pfam; PF01542; HCV_core; 1.

Pfam; PF01560; HCV_env.

LAST HCV_CORE; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01500; HCV_NS2; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS5s; 1.
Pfam; PF010271; helicase_C: 1.
Pfam; PF00998; Viral_RGRP; 1.
Probm; P186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    78.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SM00487; DEXDC;
                                     IPR002531; H
IPR002518; H
IPR004109; H
IPR000745; H
IPR001490; H
IPR002521;
IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Sim
Matches 160;
```

S

```
KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
                                                                                                                                          ; IPR002868; HCV_NS5a.; IPR002165; HCV_RGRP.; IPR007095; RNA_DOl_DS_PS.; IPR007094; RNA_DOl_PSvir.)01543; HCV_Capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                326573 MW;
                                                                                                                                                                                                                                                                                                     Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01020; HCV_NS5a; 1.
Pfam; PF00271; helicase_C: 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                           HCV_NS3.
HCV_NS4a.
HCV_NS4b.
                                         HCV_env.
                                                        HCV_NS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81.1%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 77.0
Matches 157; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00487; DEXDC;
                                                                                                            IPR000745;
IPR001490;
     IPR002522;
                      IPR002521;
                                                                        IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3010 AA;
                                                                                                                                                                                                                                                                   PF01560; PF01538; P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           D-structure.
                                                                                         InterPro;
InterPro;
InterPro;
                                                                                                                                                interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
NP_BIND
SITE
                                       InterPro;
                                                                        InterPro;
                                                                                                                                                                InterPro;
                                                                                                                                                                                 nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                        InterPro;
                                                                                                                                                                                                    nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INIT MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LRANSMEM
                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                       Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                      1125 CGSSDLYLVTRHADVIPVRRKGDSRGSLLSPRPISYLKGSSGGPLLCPSGHVVGIFRAAV 1184
                                                                        113
                                                                                                                                            CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                       01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
6enome polyprotein [Contains: dapsid protein C (Core protein) (P22);
6nvelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NG1); Protein P7; NonStructural protein NS2 (P21)
(EC 3.4.22.-); Protessc/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P65); Nonstructural protein NS5A (P65); CP7); Nonstructural protein NS5A (P65); CP7); Nonstructural protein NS5A (P65); Nonstructu
                                                                                           SEQUENCE FROM N.A.
MEDLINE-92295714; PubMed-1318627;
MEDLINE-92295714; PubMed-1318627;
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Nakazawa T., Hijikata M., Ishimura Y., Shimotchno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese carrier: sequence variation within the same individual and among infected individuals.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VITUE RES. 23:99-53(1992).

-!- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

-!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 postition, Cys or Thr in Pl and Ser or Ala in Pl'.

-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                        ATQTFLATSINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       <
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (RNA)(N).
SUBDINT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLICOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                                                                                                                                                                                  174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; D11168; BAA01943.1; -. PIR; A45573; A45573.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                           STANDARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1A10; 25-MAR-98.
1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=31642;
                                                                                                                                                                                                                                                                                                                                         POLG_HCVJT
                                                                                                                                            114
                                                                                                                                                                                                                                                                                                                          POLG_RCVJT
                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                               ò
                                                                                                                                          οý
                                     a
                                                                        ò
                                                                                                     용
```

```
REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NAJOR ENVELOPE PROTEIN NSI/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9; Gaps
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     94A1C77435D642BB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 827.5; DB 1
                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20; Mismatches
```

```
ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                  the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
MEDILINE-97015088; PubMed-8861916;
LOVE R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
The crystal structure of hepatitis C virus NS3 proteinase reveals a
trypsin-like fold and a structural zinc binding site.";
                                                                                                                                                                                                                                                                                                                                                           Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onishi E., Andoh T., Yoshida I., Okayama H.; Structure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691.
MEDLINE-98227846; PubMed-9568891;
Yan Y., Li Y., Munshi S., Sardana V., Cole J.L., Sardana M.,
Steinkuehler C., Tomei L., de Francesco R., Kuo L.C., Chen Z.;
"Complex of NS3 protease and N84A peptide of BK strain hepatitis C
virus: a 2.2-A resolution structure in a hexagonal crystal form.";
Protein Sci. 7:837-847(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4B AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
-I- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the vip precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.
-I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate
                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-96235224; PubMed-8647104;
BOCOWSKI P., Helland M., Oehlmann K., Becker B., Korneteky L.;
Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
                                                                                                                                                                          Ş
                                                                                                    Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE-91140698; PubMed-1847440;
                                                                                                                                                                                            01-AUG-1992 (Rel. 23, Created)

    Virol. 65:1105-1113(1991).

                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                        from human carriers.
                                                                                                                                                                                                                                                                                                                     WCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        [RNA] (N).
                                                                                                                                                                        POLG_HCVBK
P26663;
                            24
                                                                                                                                                      RESULT 5
                                                                                                                                                                          a
                                                               õ
                                                                                                       ò
                            ò
                                             a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its mes by non-profil institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CELLULAR AMINOPEPIDAGE.

CAPSID PROTEIN C (POTENTIAL).
MAJOR ENVELOR (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/ZE (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/ZE (POTENTIAL).
PROTEASE/HELICASE NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/ZE (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/ZE (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/ZE (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI/ZE (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
          LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA.
                                                                             SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_Psvir.
Pfam, PF0143; HCV_capsid; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01542; HCV_core; 1.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF0158; HCV_NS2; 1.
Pfam; PF01504; HCV_NS3; 1.
Pfam; PF01005; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4b; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_NS5a.
                                                                                                                                                                                                                                                                    EMBL; M58335; AAA72945.1; -. PIR; A38465; GNWVTC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        383
729
729
1006
1006
2013
3010
369
1107
11165
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                        1A10: 25-MAR 98.

1JXP: 14-JAN-98.

1C2P: 15-NOV-00.

1CSJ: 08-NOV-99.

1GX5: 09-APR-02.

1GX6: 10-APR-02.

1GX6: 10-APR-02.

1GW6: 26-JUN-00.

80HM: 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR002519;
IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              730
1007
1616
1863
2014
347
347
11083
1107
11165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         U39.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TRANSMEM
ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                 PDB;
PDB;
PDB;
PDB;
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                             53
                                                                                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                       54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                        (POTENTIAL)
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                               (POTENTIAL). (POTENTIAL).
                                                                                                           (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                6
                                                   (POTENTIAL)
                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                   Length 3010;
                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                       327189 MW; F8422D5ECCFDFD9C CRC64;
                                                                                                                                                                                                                                                                                                                                                   80.7%; Score 823.5; DB 1;
76.5%; Pred. No. 2.1e-69;
Live 21; Mismatches 18;
                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                           (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                               (GLCNAC.
                                                   (GLCNAC.
                                                                                              (GLCNAC.
              GLCNAC
                   (GLCNAC
                           GLCNAC
                                                                                                           (GLCNAC
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                 1186
1188
1197
1202
1204
                                                                                                                                                1063
1074
1076
1081
1087
1092
1094
1103
                                                                                                                                                                                                                                                     11144
11157
11161
11163
11166
11171
                                                                                         645
2041
2077
2240
2529
2788
1035
1047
                                                                                                                                                                                                                                                                                                        1187 118
1189 119
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                                                                                                                          al Similarity
156; Conserv
                                                                                                                                                                                                                                                                  1158
1162
1165
1168
1172
1175
                                                                                                                                                             1075
1077
1082
1086
1090
                                                                                                                                                                                                    1095
                                                                                                                                                                                                         1101
                                CARBOHYD
CARBOHYD
                                                         CARBOHYD
                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                                        CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                       SEQUENCE
                                                                                                                                                                                                                                                                                                                                                    Query Match
      CARBOHYD
                   CARBOHYD
                           CARBOHYD
                                             CARBOHYD
                                                   CARBOHYD
                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                           Local
                                                                                                                                                                         HELIX
TURN
STRAND
TURN
                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                     STRAND
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                     STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                               TURN
STRAND
TURN
STRAND
                                                                                                                                    HELIX
STRAND
STRAND
STRAND
                                                                                                                                                             TURN
                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                              STRAND
                                                                                                                                                                                                    STRAND
                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                            Best Loca
Matches
                                                                                                                                                                                                                                               TURN
                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                           a
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration
                                    -!- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in Pl and Ser or Ala in Pl'.
-!- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -1-SUBDAT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                               DISCUSSION OF SEQUENCE.

MEDLINE-91192160; PubMed-1849488;

Kato N., Hijkata M., Nakagawa M., Ootsuyama Y., Muraiso K.,

Ohkoshi S., Shimotohno K.;

"Molecular structure of the Japanese hepatitis C viral genome.";

FEBS Left. 280:325-328(1991).

-I- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4B ARE

HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
                                                                                                                                                                                                                                                              MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S., Sugiamra T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from Japanese patients with non-A, non-B hepatitis.";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       · I - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
              PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV_capsid.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS3.
HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_core
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; D90208; BAA14233.1; -.
              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002868;
IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; A39253; GNWVCJ
HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                             Hepacivirus.
NCBL_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [RNA](N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                        P26662;
POLG_HCVJA
```

æ

```
1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    53
                                                                                                                                                                                                                                                                                                                                                                                                                                                              CELLULAR AMINOPEPTIONSE.
CAPSID PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRINLSGD-----TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Hallacase; AFP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease INTI_MET 1 REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MW; AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                80.7%; Score 823.5; DB 1;
75.5%; Pred. No. 2.1e-69;
tive 23; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GECNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
erPro: IPR001650; Helicase_C.

LerPro: IPR007095; RNA_pol_DS_PS.

terPro: IPR007094; RNA_pol_DS_PS.

tam; PP01543; HCV_capsid; 1.

tam; PF01543; HCV_core; 1.

yfam; PF01559; HCV_env; 1.

Pfam; PF01560; HCV_NS1; 1.

Pfam; PF01560; HCV_NS1; 1.

Pfam; PF01006; HCV_NS4; 1.

Pfam; PF01006; HCV_NS4; 1.

Pfam; PF01001; HCV_NS4; 1.

Pfam; PF01001; HCV_NS4; 1.

Pfam; PF00271; helicase_C; 1.

Pfam; PF00271; helicase_C; 1.

Pfam; PF0098; Viral_RdRP; 1.

A Probom; PD186062; HCV_NS1; 1.

A SMARY; SM0487; DEXDC; 1.

A Norotein; Glycoprotein; Tran

~ Ivan profession for the protein; Env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity 75.5
Matches 154; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           645
2041
2077
2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      645 64
2041 204
2077 207
2240 224
2788 278
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11107
11165
11230
1316
196
209
234
250
305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMED cutstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                      01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 24). Last sequence update)
01-AUG-1992 (Rel. 24). Last annotation update)
Genome polyprotein [Contains Capsid protein C (Core protein) [P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.21.-); Proteinse/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.-); Proteinse/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.-); Protein Protein NS5A (P6); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P6); Nonstructural protein
NS5B (P66) (P70) (RM-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate HC-J8) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flavivridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genotypes.;
Virology 188:331-341(1992).
VIROLOGESTHE MEMBRANE-RELATED FUNCTION.
VIROLOGESTHE NOTE IN THE VIRAL RNA REPLICATION.
VIROLOGY AND THE VIRAL RNA REPLICATION.
VIROLOGY OF COUNCIL IN THE VIRAL RNA REPLICATION.
VIROLOGY OF COUNCIL IN THE VIRAL RNA PROPERTY OF COUNCIL OF VIRAL RNA PROPERTY OF CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [RNA](N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
MEDLINE-92230232; PubMed-1314459;
Okamoto H., Kurah R., Okada S., I., Yamamoto K., Lizuka H., Tanaka 'Okamoto B., Tsuda F., Mishiro S.;
Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: comparative study of four distinct
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                           3033 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid.
HCV_core.
HCV_RS1.
HCV_NS1.
HCV_NS3.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
HSSP; P27958; 1HEI.
                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR002531; H
IPR002518; H
IPR004109; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                 Hepacivirus
                                                                                                       POLG_HCVJ8
P26661:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                          POLG_HCVJ8
                                                                                                                            셤
```

6

```
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1034 TAYTQQTRGLLGAIVVSLTGRDKNEQAGQVQVLSSVTQTFLGTSISGVLWTVYHGAGNRT 1093
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                         CELUCLER AMINOPEPTIONSE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN C (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Gaps
                                                                                                                                                                                                                                             Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; AFP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease. INTI_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ( POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ; Score 675; DB 1; Length 3033;
; Pred. No. 2.4e-55;
24; Mismatches 30; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1A173E7E3381FD1A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA; 330177 MW;
                                                              Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF01506; HCV_NS5a; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 66.2%;
69.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         st Local Similarity 69.89
tches 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3033
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   233
2299
3305
417
423
430
                    PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
SOLUTION AND CONTRACT OF THE STATE OF THE ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

PRT; 3033 AA

STANDARD;

POLG_HCVJ6 ID POLG_HCVJ6 AC P26660;

RESULT 8

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified annothed in statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.

MEDLINE-92044440; PubMed-1658196;
A MEDLINE-92044440; PubMed-1658196;
A MEDLINE-92044440; PubMed-1658196;
A Machida A., Miyakawa Y., Mayumi M.;
Mucleotide sequence of the genomic RNA of hepatitis C virus isolated
T from a human carrier: comparison with reported isolates for conserved
and divergent regions.;
J. Gen. Virol. 72:2697-2704(1991).
J. Gen. Virol. 72:2697-2704(1991).
J. Gen. Virol. 72:2697-2704(1991).
C --- FUNCTION: THE SNALL PROPER MEMBRANE-RELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
C --- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.
C --- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 21, Last annotation update)
03-8-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GPN0) (NRI); Proteate P7: Nonstructural protein NS2 (GP3) (GP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INTERPRO) IPRO0252; HCV_capsid.
INTERPRO; IPRO0252; HCV_capsid.
INTERPRO; IPRO02519; HCV_env.
INTERPRO; IPRO02519; HCV_NS1.
INTERPRO; IPRO04519; HCV_NS2.
INTERPRO; IPRO04045; HCV_NS3.
INTERPRO; IPRO0445; HCV_NS4.
INTERPRO; IPRO0145; HCV_NS4.
INTERPRO; IPRO0145; HCV_NS5.
INTERPRO; IPRO02668; HCV_NS5.
INTERPRO; IPRO01669; HCV_NS5.
INTERPRO; IPRO01669; HCV_NS5.
INTERPRO; IPRO01669; HCV_NS5.
INTERPRO; IPRO01695; HCV_NS5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; D00944; BAA00792.1; -. PIR; JQ1303; JQ1303.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_core; ]
HCV_env; 1.
HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS4b;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1HEI.
                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSSP; P27958; 1HE
MEROPS; S29.001;
MEROPS; U39.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01001;
PF01506;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [RNA] (N).
                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
```

```
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                      n: Envelope protein; Helicase; ATP-binding;
ral protein; Hydrolase; Serine protease,
RENVED FROM CAPEID FNOTEIN C BY THE
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
PROTEASE/HELICASE NSI (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
Pfam: PF00271; helicase_C; 1.
Propose P00999; Viral_RGRP; 1.
Prodom; PD186062; HCV_NS1; 1.
SNART; SN00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Tocke protein; Holicase; APP-binding; Transmenrane; Nonstructural protein; Hydrolase; Serine proteing: INII_MET
                                                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   .;
0
                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                       POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                           (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F957F5C1A273BE9E CRC64;
                                                                                                                                                                                                                               (BY
(BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HHOA_ARATH STANDARD; PKT; 321 AA. 09SEL7; 049507; 16-0CT-2001 (Rel. 40, Created) 16-0CT-2001 (Rel. 40, Last sequence update) 15-SEP-2003 (Rel. 42, Last annotation update) Protease HhoA, chloroplast precursor (EC 3.4.21.-). HHOA OR AT4G18370 OR F28J12.30.
                                                                                                                                                                                                                             CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          66.0%; Score 673; DB 1;
68.7%; Pred. No. 3.7e-55;
ivc. 27; Mismatches 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
                                                                                                                                                                                                                                                                 ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                      N.LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   329165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity ....
                                                                                                                                                                    1619
1866
2017
3033
3033
369
1087
1111
1169
11241
1323
196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ź
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2811
3033
                                                                                                                                                                                                                TRANSMEM
ACT_SITE
                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HUOA_ARATH
TD HUOA_AR
TO HUOA_AR
TO 16-0CT
DT 16-0CT
DT 15-SEP
DE PEOTCESS
OG NHOA OF
OG SATABIE
CO SPETABLE
CO SPETABLE
                                                                                                                      CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                          CHAIN
                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

Arabidopsis thaliana (Mouse-ear cress). Eukaryota: Viridiplantae: Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

```
RE STRAIN-CY. Columnia.

RE STRAIN-CY. Columnia.

RA MAYER E. No. Schueller C., Wambutt R., Murphy G., Volckaert G., RA Mayer K.F. Schueller C., Wambutt R., Murphy G., Volckaert G., RA Mayer K.F. Schueller C., Wambutt R., Machar M., Entian R., D., Terryn N., RA Mayer K.F. Schueller C., Wambutt R., Machar B., Machar B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                 Lensch M.H.A., Sokolenko A., Herrmann R.G.;
"Identification and characterization of the chloroplast HhoA protease,
a homolog to the bacterial periplasmic protease HhoA.";
Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             shown due to erroneous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAUTION: Ref. 2 sequences differ from that shown due to erroneougene model prediction. AT4G18370 and AT4G18375 were originally fused into a single gene.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "The chloropiast lumen from Arabidopsis thaliana.";
Submitted (JUL-2001) to the SWISS-PROT data bank.
-!- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE OF 72-82; 96-110; 150-159; 178-211 AND 306-320. Schubert M., Peterson U., Funk C., Haas B., Schroeder W.P., Kieselbach T.;
eurosids II; Brassicales; Brassicaceae; Arabidopsis;
NCBL_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nature 402:769-777(1999).
                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                       [2]
SEQUENCE FROM N.A.
```

```
138
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ģ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12;
                modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182 --DPDNDLAVLKIETEGRELNPVVLGTSNDLRVGQSCFAI----GNPYGYENTLTIGVV 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140 ----SLLSPRPISYLK-----GSSGGPLLCPAGHAVGIFRAAVSTR--GVAKAV 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          235 SGLGREIPSPNGKSISEAIQTDADINSGNSGPLLDSYGHTIGVNTATFTRKGSGMSSGV 294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           95 KTSPSVVYIEAIELPKTSSGDIL------127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               128 ----TGSGFVWDKLGHIVTNYHVIAKLATDQFGLQRCKVSLVDAKGTR--FSKEGKIVGL 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57
non-profit institutions as long as its content is in no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58 FLATSINGVLW-----TVYH----TVYH-----GAGTRIIASPKGPVTQM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
STORAIN-ATCC 15692, PAO1;
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P., Hickey M.J., Brinkman F.S.L., Hufnagle M.O., Kowalik D.J., Lagrou M., Garber R.L., Goltry L., Tolentino E., Westbrock Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Smith M.A., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 KKKGSVVIVGRINL ---- SGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            90 YTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRG------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score v. 0.66;
Pred. No. 0.66;
"wiematches 68; Indels 100; Gaps
                                                                                                                                                                                                                                                                                   Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide.
TRANSIT 27 71 THYLAKOID.
CHAIN 72 321 PROTEASE HHOA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
                                                                                                                                                                                                                                                                                                                                                                              POLY-GLU.
CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
CHARGE RELAY SYSTEM (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Match 8.5%; Score 87; DB 1; Length 321; Local Similarity 21.3%; Pred. No. 0.66; les 54; Conservative 31; Mismatches 68; Indels 31; Mismatches 68; Mismatches 68; Mismatches 68; Mismatches 68; Mismatches 68; Mismatches 68; Misma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   209 A.A.
                                                                                                         EMBL; AF114386; AAF24060.1; --
EMBL; AL021710; CAA16717.1; ALT_SEQ.
EMBL; AL161548; CAB78839.1; ALT_SEQ.
MEROPS; S01.279; --
InterPro; IPR001940; Protease2C.
InterPro; IPR001244; Ser_protease_Try.
Fram; PF00089; trypsin; 1
PRINTS; PR00834; PROTEASES2C.
                                                                                                                                                                                                                                                                                                                                                                                                   145 145 CH
186 186 CH
264 264 CH
40 40 RM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           183 DF-IPVESLETTM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEALPIDIVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                   NIWOO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29HX08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PA4019.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          92 NVDKDLVGWQAPQGSRSLTP-----CTCGSSDL-----YLVTRHADVIPVRRRGDS 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RGSLLSPR--PIS----YLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       43 QVEGEVQ-IVSTATQTFLATSINGVL------WTVYHGAGTRTIASPKGPVTQMYT 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29 QEEREVHFLISKAAQLVMATETDVALPAKPQAMQAFLTEYCGAAAGQI------RVFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEGI_ARATH STANDARD; PRT; 437 AA.

022609; OBLK85;

16-CCT-2001 (Rel. 40, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

28-FEB-2003 (Rel. 41, Last annotation update)

Protectase Do-like 1, chloroplast precursor (EC 3.4.21.-).

Protectase Do-like 1, chloroplast precursor (EC 3.4.21.-).

Arabidopsis thaliana (Mouse-ear cress).

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
MEDIINE-20363099; PubMed-10907853;
Raneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
"Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
and BAC clones.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEGUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
Ildentification and characterization of Degp, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
Nature 406:959-964(2000).
-i- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE /
PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR003382; Flavoprotein.
Pfam; PF02441; Flavoprotein; 1.
Hypothetical protein; Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 WW; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1; Length 209;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 8.4%; Score 85.5; DB 27.9%; Pred. No. 0.54; ative 16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AE004818; AAG07406.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DNA Res. 7:217-221(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; H83144; H83144.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     [3]
SEQUENCE OF 104-118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                190 LET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   187 LNT 189
```

```
AAMP HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NIWO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loca
Matches
             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMEL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       70 VYHGAGTRIIASPKGPVTQMY------QA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA
         Gaps
                                                                                                                                                                                                                                                                                                                                                                                               CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
HSPPSSQLSNST -> SSTLFLHSPPSSHL (IN REF.
                                                                                                                                                                                                                                                                                                                                       Transit peptide; Chloroplast; Thylakoid.
CHLOROPLAST (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SYL------KGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -1- SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               / Match 8.2%; Score 83.5; DB 1; Length 437; Local Similarity 26.2%; Pred. No. 2; nes 45; Conservative 17; Mismatches 55; Indels 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1497BlAB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                     2).

V -> I (IN REF. 2).

P -> S (IN REF. 2).

G -> R (IN REF. 2).

G -> D (IN REF. 2).

L, -> HF (IN REF. 2).

L -> V (IN REF. 2).

I -> V (IN REF. 2).

I -> V (IN REF. 2).

O -> E (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                  PROTEASE DO-LIKE 1. SERINE PROTEASE.
                                                                                                 -1- INDICTION: By heat shock.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
-1- SIMILARITY: Contains 1 PD2/DHR domain.
                                                                                                                                                                                                                                              MEROPS; S01.279; --
InterPro; IPR001478; PDZ.
InterPro; IPR001478; PDZ.
InterPro; IPR001240; Protease2C.
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00089; trypsin; 1.
Pfam; PR0089; trypsin; 1.
Primrs; PR0084; PR0TEASESZC.
SWART; SM00228; PDZ; 1.
PROSITE; PS50106; PDZ; 1.
                                                                                                                                                                                                                                                                                                                                                           THYLAKOID.
                                                                                                                                                                                                                    EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              46213 MW;
                                                                                                                                                                                                                                                                                                                                      Aydrolase; Serine protease; IRANSIT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                360
60
60
80
138
138
14
  Columbia;
                                                                       O-PHENANTHROLINE
                                                                                                                                                                                                                                                                                                                                                                   104
152
324
                                                                                           MEMBRANE
                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
ACT_SITE
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                          TRANSIT
CHAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

RESULT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL cutstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  197 WMEWH------PRAPVLLAGT-ADGNTWMWKVPNGDCKTFQGPNCPATGGR----- 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      123 TRHADVIPVRRR---GDSRGS-----LLSPRPISYLKGSSG--GPLLCPA------ 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   68 WIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWQAPQGSRSL----IPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241 -----VLPDGKRAVVGYEDGTIRIWDLKQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                                                                                                                                              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HEPARIN-BINDING (POTENTIAL).
POLY-GLU.
WD 1.
WD 2.
WD 3.
WD 4.
WD 5.
WD 6.
WD 7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ------GHAVGIFR----AAVSTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8.1%; Score 83; DB 1; Length 452; 25.3%; Pred. No. 2.4; Live 13; Mismatches 47; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GO; GO:0008201; F:heparin binding activity; TAS.
InterPro; IPR001680; WH40.
Pfam; PF00400; WD40; 8.
SMART; SM00320; WD40; 8.
PROSITE; PS00678; WD_REPEATS_1; 1.
PROSITE; PS50082; WD_REPEATS_2; 6.
PROSITE; PS50294; WD_REPEATS_26.
                                            15-JUL-1998 (Rel. 36, Created)
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
452 AA
                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
PRT;
                                                                                                                                                                                                                                                                                                                       TISSUE-Brain;
MEDLINE-95262124; PubMed-7743515;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           = = -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M95627; AAA68889.1; -. PIR; I39383; I39383.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genew; HGNC:18; AAMP.
                                                                                                                                                                        Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  452 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Repeat; WD repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
les 42; Conserv
                                                                                                                                                                                                                                               NCBI_TaxID-9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              603488;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
```

225 DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGGG 258

296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341

g

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         109 GTGVYKHCVWGAGSSSTGTTASPSTETCSQHAT----LVGGTSKPFWLVPGGTGNNGNCG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           165 CGGGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGGGSSSSSSCIHIKVENTDEQFL 224
                                                                                                                                                                                                                                               Praser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R. Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren R., Gwinn M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Fichardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SGDTAYA - - - - - QQTRGEQGCQKTSH - - - - TGRDKNQVEGEVQIVSTATQTFLATSI -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVMSIATDGNTFVLACVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       64. -NGVLWTVYHGAG---TRTIASPKGPVTQMYTNVDKDLVG-----WQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           114 C----GSSDLYLVTRHADVIP-----VRRRGDSRGSLLSPRPISYLK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                     Science 281:375-388(1998).
                                                                                                                                                  Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL.
HYPOTHSTICAL LIPOPROTEIN TP0136.
N-ACYL DIGLYCERIDE (POTENTIAL).
GLY/SER-RICH.
GLY/SER-RICH.
GLY/SER-RICH.
POLY-SER.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                   "Complete genome sequence of Treponema pallidum, the syphilis spirochete.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                         -1- SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 1; Length 485;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   77; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TIGR: TP0136; -.
Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---GSSGGPLLCPAGHAVG 167
                                              16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 6.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 78.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; AE001199; AAC65137.1; ALT_INIT.
                                                                                                                                                                                                                    STRAIN-Nichols;
MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 48984 NW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Complete proteome.
SIGNAL 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                  Treponema pallidum
                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                         (Potential).
                                                                                                                                                                  NCBI_TaxID=160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   50;
Y136_TREPA
ID Y136_TREPA
AC 083172;
                                                                                                                                                    Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LSC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LIPID
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                 셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
this SWISS-PROT entry is copyright. It is produced through a collaboration the burden the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its meb by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                        SEQUENCE FROM N.A.

MEDLINE-94102535; PubMed-7503814;

MEDGINE-94102535; DubMed-7503814;

Ramos R.G., Igloid G.L., Lichte B., Baumann U., Maier D.,

Schneider T., Brandstaetter J.H., Froehlich A., Fischbach K.-F.;

"The irregular chiasm C-roughest locus of Drosophila, which affects
aronal projections and programmed cell death, encodes a novel
immunoglobulin-like procein.";

Genes Dev. 7:2533-2547(1993).
                                                                                                                                                                                                                                                                                                                                                                                       -:- SUBCELLULAR LOCATION: Type I membrane protein.
-:- TISSUE SPECIFICITY: POSTEMBRYONIC EXPRESSION IS STRONG IN THE DEVELOPING OPTIC LOBE AND IN THE EYE IMAGINAL DISC.
-:- DEVELOPMENTAL STAGE: STRONGLY EXPRESSED IN EMBRYOS. ALSO FOUND IN LATE LARVAL AND PUPAL STAGES.
-:- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     domain; Glycoprotein; Signal; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IGLIKE C2-TYPE 3.
IG-LIKE C2-TYPE 4.
IG-LIKE C2-TYPE 5.
GLN-RICH (OPA-REPEAT).
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                          -!- FUNCTION: REQUIRED FOR CORRECT AXONAL PATHWAY FORMATION IN THE OPTIC LOBE AND FOR PROGRAMMED CELL DEATH IN THE DEVELOPING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL.
IRREGULAR CHIASM C-ROUGHEST PROTEIN.
EXTRACELLULAR (POTENTIAL).
                                                                            Irregular chiasm C-roughest protein precursor (IRREC protein).
                                                                                                Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Beoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CYTOPLASMIC (POTENTIAL).
IG-LIKE C2-TYPE 1.
IG-LIKE C2-TYPE 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; 221641; CAA79756.1; -.
EMBL; L11040; AAA16632.1; -.
PIR: A49448 A49448.
FINS A9448 A49448.
FINS EMBLS A1948.
InterPro; IPR007110; Ig-like.
InterPro; IPR007110; Ig-like.
InterPro; IPR003006; Ig_MRC.
Pfam; PF00047; Ig: 4.
SMART; SM00408; IG-2; I.
PROSITE: PSSO893; IG_LIKE; 5.
Transmembrane; Immunoglobulin domain; Glycoprotein;
                           01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
764 A.A.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19
5533
5533
1123
123
123
143
143
153
16
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Cell adhesion SIGNAL
CCR_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
DOMAIN
TRANSMEM
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          OMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
```

```
FF
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                          7:
                                                                                                                                                                                                                                                                      109 --LTPCTCGS--SDLYLVTRHADVIPVRRRGD----SRGSLLSPRPISYLKGSSGGPLLC 160
                                                                                                                                                              504 GKYNCTVVNDYGND----VAEIQLQAKKSVSLLMTIVGGISVVAFLLVLTILVVVYIKCK 559
                                                                                                                                                                                                                                  560 KRTKLPPADVISEHQITKNGGVSCKLEPGDRTSNYSDLKVDISGGYVPYGDYSTHYSPPP 619
                                                                                                                                                                                                                                                                                                        620 QYLTTCSTKSNGSSTIMQNNHQNQLQLQQQQQQSHHQHHTQTTTLPMTFLTNSSGSL-- 677
                                                                                                                          27 GEQGCQKTSHTGRDKNQVEGEVQIVSTATQIFLATSING------VLWTVYHGAG 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Delassus S., Sonigo P., Wain-Hobson S.;
"Genetic organization of 41bbon ape leukemia virus.";
Virology 173:205-213(1989)
--- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
                                                                                                                                                                                                  76 TRIIASP--------KGPVTQMYINVDKDLVGWQAPQGSRS-----
                                                                                            Gaps

    -!- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
    -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2.

                                                                                        53;
507 507 N-LINKED (GLCNAC. .) (POTENTIAL) 764 AA: 82947 MM; 262225D2B2AlC181 CRC64;
                                                    Length 764;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-MAY-1991 (Rel. 18, Created)
01-KKY-1991 (Rel. 18, Last sequence update)
01-KKY-1991 (Rel. 41, Last annotation update)
Pol polyprotein [Contains: Protease (EC 3 4.23.-); Reverse transcriptase (EC 2.7.7.49); Endonuclease).
                                                  7.7%; Score 78.5; DB 1; Length 76
20.3%; Pred. No. 12;
tive 27; Mismatches 69; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; Retroid viruses; Retroviridae; Gammaretrovirus NCBL_TaxID=11840;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE; PS00141; ASP_PROTEASE; 1.
PROSITE; PS50175; ASP_PROT_RETROV; 1.
Hydrolase; Transferase; RNA-directed DNA polymerase;
Aspartyl protease; Endonuclease; Polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 1165 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001995; Aspprotease_rtrv.
InterPro; IPR001969; Aspprotease_site.
InterPro; IPR002156; RNaseH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SECTENCE FROM N.A.
MEDLINE-90051069; Pubmed-2683360;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M26927; AAA46810.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR001584; Rve.
InterPro; IPR000477; RVTse.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gibbon ape leukemia virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00075; rnaseH; l. Pfam; PF00655; rve; l. Pfam; PF00077; rve; l. Pfam; PF00078; rvt; l.
                                                      Ouery Match 7.78

Best Local Similarity 20.38

Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                  | :|
678 -TGSIIG 683
                                                                                                                                                                                                                                                                                                                                           161 PAGHAVG 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GNLJGL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; B32595; GNLJGL
HSSP; P03355; 1MML.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEROPS; A02.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 {DNA}(N).
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POL_GALV
                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 15
POL_GALV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        £ 8
                                                                                                                                                                                                                                a
                                                                                                                                                        a
                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                            g
```

```
411 PITPROVREFIGIAGECRLWI----PGFASLAAPLYPLIKESIPFIWTEEHQQAFDHIKK 466
                                                                                                                                                                                                                                                                                                                                                                                                                               96 DLVGWQA---PQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSL---LSP--RPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  467 ALLSAPALALPDLTKPFT-----LYI-----DERAGVARGVLTQTLGPWRRPV 509
                                                                                                                                                                                                                     ---RDKNQVEGEV 48
                                                                                                                                                                                                                                                                                                                          49 QIVSTATQTFLATSINGVLWTVYHGAGTRTIASPKGPVT-------OMYTNVDK 95
                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          148 SYLKG----SSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64;
                                                                                                          Query Match 7.6%; Score 78; DB 1; Length 1165; Best Local Similarity 22.6%; Pred. No. 22; Matches 53; Conservative 33; Mismatches 85; Indels
                            27 27 BY SIMILARITY.
1165 AA; 129886 MW; 8B7AFD54812B7E1A CRC64;
                                                                                                                                                                                                                     3 KKGSVVIVGRINLSGDIAYAQQIRGEQGCQK-TSHTG-
PROTEASE
                            ACT_SITE
SEQUENCE
     CHAIN
```

Search completed: August 30, 2003, 19:13:49 Job time : 10.7567 secs

Sequence:

Run on:

Searched:

```
091748 hepatitis of 091748 hepatitis of 091776 hepatitis of 091776 hepatitis of 09178 hepatitis of 00178 hep
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
MEDLINE-20323484; PubMed-10864644;
Lal V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Lal V.C., Anong W., Skelton A., Ingravallo P., Vassilev V.,
Bonis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus.";
[2]
Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Genome polyprotein.
Mucosal disease virus.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Lai V.C.H., Hong Z.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL: AF568278; AAF82566.1; ~.
HSSP; P26663; lJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                        AL IGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS; S31.001; ...
InterPro; IPR000280; CDvir_endptseP80.
InterPro; IPR001410; DEAD.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001656; HCV_RARP.
InterPro; IPR001055; Myb_DNA_binding.
InterPro; IPR001055; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF002971; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
091R81
091R08
091R00
036609
091R86
091R83
091R74
091R73
091R75
091R75
091R75
091R76
091R76
091R77
091R77
091R70
                                                                                                                                                                                                                                                                                                                          Q91RR9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-OCT-2000 (TrEMBLrel. 15, 01-OCT-2000 (TrEMBLrel. 15, 01-MAR-2003 (TrEMBLrel. 23,
 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11099;
Pestivirus.
                        837
836.5
836.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9IFH8
                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 1
Q9IFH8
Q9ifhB mucosal dis
O3679 hepatitis c
O81756 hepatitis c
O9165 hepatitis c
O3608 hepatitis c
O3608 hepatitis c
O99wx5 hepatitis c
O91rr8 hepatitis c
O91rr5 hepatitis c
                                                                                 August 30, 2003, 19:00:22; Search time 37.5921 Seconds (without alignments) 1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                          1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Description
         GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                        830525 segs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                           - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              036579
081756
091FE5
09ELS8
003463
03608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   09PWU9
091RR8
091RR5
091RR5
091RR9
091RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  091FH8
                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                       sp_archea:*
sp_bacteria:*
sp_fung::*
sp_human:*
sp_invertebrate:*
sp_mammal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sp_rodent:*
                                                                                                                                 US-09-965-594-20
1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_plant:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 sp_virus:*
                                                                                                                                                                                                                                                                                                                                                              SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4040
3011
2436
3011
3011
3011
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
                                                            OM protein
                                                                                                                                                                                                                                                                                                                                                              Database :
```

Result

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   081756
081756;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DON KEEN DON WAS SON W
                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                          Ä
                                                                                                                                                                                                                                                                                                                                                                     122 VTRHADVIPVRRGDSKGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                              10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase. SEQUENCE 4040 AA; 453073 MW; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                            5 GSVVIVGRINLSGD---TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                        3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-97373636; Pubmed-9228008;
Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone
                                                                                                                                                                                                      DB 12; Length 4040;
                                                                                                                                                                                                  Ouery Match 85.5%; Score 872.5; DB 12; Length Best Local Similarity 89.2%; Pred. No. 1.4e-74; Matches 174; Conscrvative 5; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3011 AA.
PRINTS; PR00729; CDVENDOPTASE.
SMART; SM00487; DEXDc; 1.
SMART; SM00440: HELICC; 1.
PROSITE; PS00037; MYB_1; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50501; RDRP_VIRAL; 1.
PROSITE; PS00531; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Helicase_C.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  182 VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       036579;
01-JAN-1998 (TrEMBLrel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ....erPro; IPRO02519; HC
InterPro; IPRO02519; HC
InterPro; IPRO02518; HC
InterPro; IPRO041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Rice C.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-H77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

```
54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN -----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                               9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    P SEQUENCE FROM N.A.

Choo O. L., Richman K., Han J.;
The nuclectide sequence of the Hepatitis C viral genome.*;
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.

EMBL: M32084; AAA45677.1;
R InterPro: IPR001410; DRAD.

R InterPro: IPR001531; HCV_NS1.
R InterPro: IPR001538; HCV_NS3.
R InterPro: IPR001499; HCV_NS4a.
R InterPro: IPR001499; HCV_NS4b.
R InterPro: IPR001496; HCV_NS4b.
R InterPro: IPR001509; HCV_NS4b.
R InterPro: IPR001650; Helicase_C.
R InterPro: IPR007095; RNA_POl_DS_PS.
R InterPro: IPR007099; RNA_POl_DS_PS.
R InterPro: IPR007094; RNA_POl_DS_PS.
R InterPro: IPR007094; RNA_POl_DS_PS.
R Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; Pred. No. 2.2e-73;
10; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (TrEMBLrel. 01, Created)
(TrEMBLrel. 01, Last sequence update)
(TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 2436 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 STRGVAKAVDFIPVESLETTMRSP 197
                                     Pfam; PF01539; HCV_env; 1.
Pfam; PF01560; HCV_env; 1.
Pfam; PF01560; HCV_NS; 1.
Pfam; PF02907; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF010271; Hclicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein (Fragment). Hepatitis C virus.
HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 82.88
Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TrEMBLrel
01-NOV-1996 (TrEMBLrel
01-MAR-2003 (TrEMBLrel
```

ä

```
Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9ELS8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q9ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 5
Q9ELS8
                DOR KEEN DOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23
DR Pfan: PF01006; HCV_NS4a: 1.

DR Pfan: PF01006; HCV_NS4a: 1.

DR Pfan: PF01506; HCV_NS5a: 1.

DR Pfan: PF00271; Hclicase_C: 1.

DR Pfan: PF00271; hclicase_C: 1.

DR ProDom: PD186062; HCV_NS1: 1.

DR PROSITE: PS0507; RDRP_POSITIVE; 1.

DR PROSITE: PS050507; RDRP_VRAL: 1.

TW ATP-binding: Coat protein: Envelope protein; Glycoprotein; W Hydrolase: Nonstructural protein: Polyprotein:

TW NNA directed RNA polymerase: Transferase: Transmembrane.

TW NNA_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----I.SGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              555 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       sequence..;
J. Gen. Virol. 82:1291-1297(2001).
J. Gen. Virol. THE VIRION GRAPHED BY A LIPOPROTEIN B. THE ENVELOPE CONSISTS OF TWO PROTEINS.
PROTEIN M AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
EMBL, RE71662; AAF81759.1;
HSSP; P27958: 1AIV.
INTERPO: IPRO000345, CYLC_heme_bind.
InterPro: IPRO0003410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-21262212; PubMed-11369872;
Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.,
"Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-0CT-2000 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein
Heparitis C virus.
Heparitis C virus.
Heparitis SRNA positive-strand viruses, no DNA stage: Flaviviridae;
Heparivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Match 83.8%; Score 854.5; DB 12; Length 2436; Local Similarity 82.8%; Pred. No. 4e-73; es 169; Conservative 9; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2436 2436
2436 AA; 264734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro: IPR00252; HCV_capsid. InterPro: IPR002521; HCV_core. InterPro: IPR002521; HCV_core. InterPro: IPR002531; HCV_NS1. InterPro: IPR002518; HCV_NS2. InterPro: IPR004109; HCV_NS3. InterPro: IPR000149; HCV_NS3. InterPro: IPR001490; HCV_NS3. InterPro: IPR001490; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q9IFES
Q9IFES;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOR WEEK WAY STATE OF SOLUTION OF STATE OF SOLUTION OF STATE OF SOLUTION OF STATE OF SOLUTION OF SOLUT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
```

```
114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE; PS00190; CTOCHROME_C; 1.
PROSITE; PS00190; CTOCHROME_C; 1.
PROSITE; PS50507; RARE_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
FNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     83.8%; Score 854.5; DB 12; Length 3011; 82.8%; Pred. No. 5.3e-73; Live 9; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interpro: IPR00156; HCV_RGAP.
Interpro: IPR001550; Helicase_C.
Interpro: IPR001050; Helicase_C.
Interpro: IPR001050; Helicase_C.
Interpro: IPR001054; RNA_pol_DS_PS.
Interpro: IPR001054; RNA_pol_PSvir.
Pfam; PF01542; HCV_core; 1.
Pfam; PF01550; HCV_core; 1.
Pfam; PF01550; HCV_MS1; 1.
Pfam; PF01000; HCV_MS2; 1.
Pfam; PF01001; HCV_MS3; 1.
Pfam; PF0100998; VIRAL_RGRE; 1.
Proponer, PF06602; HCV_MS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HSSP; P27958; 1HEI.
InterPro; IPR000345; CytC_heme_bind.
InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           al Similarity 82.8
169; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-colonel;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DAY REPRESENTED THE PROPERTY OF THE PROPERTY O
```

```
MEDLINE-92044440; pubMed-1658196; Okandor H., Machida A., Okandor H., Okada S., Suglyama Y., Kurai K., Iizuka H., Machida A., Miyakawa Y., Mayumi M.; Mayumi M.; Mayumi M.; Mayumi M.; Mayumi M.; Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human cariler: comparison with reported isolates for conserved and divergent regions.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Inchauspe G.;
In
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydroclase; Nonstructural protein; Polyprotein; RNA Nonstructural protein; Polyprotein; RNA directed RNA polymerase; Transferase; Transmembrane. SEQUENCE 3011 AA; 327112 MW; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                       Dirachine-Ni-Ji; PubMed-1335573; MEDLINE-317120; PubMed-1335573; Okamoto H., Kanal N., Mishiro S.; "Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-Jl) with high homology to USA isolates."; Nucleic Acids Res. 20:6410-6410(1992).
"The 5'-terminal sequence of the hepatitis C virus genome."; Jpn. J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Okamoto H.;
Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR002166; HCV_RdRP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; I.
Pfam; PF01542; HCV_core; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE; PS50507; RDRP_POSITIVE; 1. PROSITE; PS50521; RDRP_VIRAL; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-HC-J1;
MEDLINE-94174722; PubMed-7510436;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_capsid.
HCV_core.
RCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002531; HCV_NS1.
IPR004109; HCV_NS2.
IPR004109; HCV_NS3.
IPR000745; HCV_NS4a.
IPR001490; HCV_NS4b.
IPR002868; HCV_NS5b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01538; HCV_MS2; 1. Pfam; PF01907; HCV_MS2; 1. Pfam; PF01907; HCV_MS4a; 1. Pfam; PF01001; HCV_MS4b; 1. Pfam; PF01506; HCV_MS5a; 1. Pfam; PF001506; HCV_MS5a; 1. Pfam; PF001506; HCV_MS5a; 1. Pfam; PF00150999; Viral_RGRP; 1. Probom; PD186062; HCV_MS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_env; 1
HCV_NS1; 1
HCV_NS2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEXDC:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002519;
                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-HC-J1;
                                                                                                         STRAIN-HC-J1
                                                                                                                                                                                                                                                                                                                                                                               STRAIN-HC-J1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
     ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KRGSVVIVGRIN------LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match 83.8%; Score 854.5; DB 12; Length Best Local Similarity 82.4%; Pred. No. 5.3e-73; Matches 168; Conservative 11; Mismatches 16; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS001007; RDRP_POSITIVE; 1.
PROSITE; PS0501; RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein Hydrolase; Monstructural protein; Polyprotein; Clycoprotein RNA-directed RNA polymerase; Transferase; SEQUENCE 3011 AA; 327107 MW; A6BECFSA3B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TIEMBLEEL. 01, Created)
01-NOV-1996 (TIEMBLEEL. 01, Last sequence update)
01-MAR-2003 (TIEMBLEEL. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT: 3011 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR001206; HCV_RG4b.
InterPro; IPR001050; Helicase_C.
InterPro; IPR001050; RNA_pol_DS_PS.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_core; I.
Pfam; PF01539; HCV_env: I.
Pfam; PF01589; HCV_NS1; I.
Pfam; PF01589; HCV_NS2; I.
Pfam; PF01589; HCV_NS3; I.
Pfam; PF01006; HCV_NS4s; I.
Pfam; PF01006; HCV_NS4s; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-91013116; PubMed-2170712;
  HCV_capsid
                                                                                                         IPR002518; HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEXDC;
                           IPR002521; I
IPR002519; I
IPR002531; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus
  InterPro;
                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
```

a

ò g ò q õ

SMART;

Q03463 Q03463;

RESULT 6 003463

```
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
InterPro;
                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                     174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              chimeras
                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                     09PWX5;
                                                                                                                                                                                                                                                                                                                                            Q9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
HSSP;
                                                                                                                                                                                                                                                                                                                        œ
                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                  09PWX5
                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                        g
                                                                                                                                                                                     ò
                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                      ö
                                                             54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                           Yanngi M., Purcell R.H., Emerson S.U., Bukh J.;
"Transcripts from a single full-length cDNA clone of hepatitis C virus
are infectious when directly transfected into the liver of a
                                                 3 KKGSVVIVGRIN-----LSGDTAYAQQIRGEQGCQKTSHTGRDKNQVEGEVQIVST
                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ø
                                                                                                                                                                                                                                                                                                  Genome polyprotein.
Hepatitis C virus strain H77.
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).
--- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE. CONSIGTS OF TWO PROTEINS: PROTEIN HAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: AFOLI751: AAB67036.1; ---
HSSP: P27958; IHEI.
                             6
       DB 12; Length 3011;
                             Indels
                                                                                                                                                                                                                                                                    01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                   6e-73;
                                                                                                                                                                                                                                                  PRT; 3011 AA
                             8; Mismatches
        Score 853.5;
Pred. No. 6.6
                                                                                                                                                                                       InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002519; HCV_core.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR0002518; HCV_NS2.
InterPro; IPR0004109; HCV_NS4.
InterPro; IPR000409; HCV_NS4.
InterPro; IPR001409; HCV_NS4b.
InterPro; IPR002166; HCV_NS4b.
InterPro; IPR002166; HCV_RS4b.
InterPro; IPR002166; HCV_RSP.
InterPro; IPR0010550; HRA_POl_DS_PS.
InterPro; IPR0010550; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
                                                                                                                                                                          174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-97385173; PubMed-9238047;
        83.7%;
82.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_core;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV_NS4a;
                            Matches 169; Conservative
                                                                                                                                                                                                                                                  PRELIMINARY;
                   Similarity
                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                  Hepacivirus.
NCBI_TaxID=63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PF01001;
PF01506;
PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01560;
PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF02907;
                                                                                                                                                                                                                                                                                                                                                                                                                             chimpanzee.";
                                                                                                                                                                                                                                                                                                                                                                           STRAIN-H77
         Query Match
                    Best Local
                                                                                                                                                                                                                                                  036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pfam;
                                                                                                                                                                                                                           g
                                                                  g
                                                                                         ö
                                                                                                           셤
                                                                                                                                 ò
                                                                                                                                                                           ò
```

```
1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BUKH J.;
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
ILPOPROTEIN THE VIRION OF THIS VIRIOS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Tran
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE-99420396; PubMed-10489358;
Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
Yanagi M., Purcell R.H., Emerson S.U., Bukh J.;
Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic la and 2a chimeras.":
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ٠<u>.</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Pred. No. le-72;
10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                   327112 MW; 0B75E6B81CB5C198 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01.WAY-2000 (TrEMBLrel. 13, Created)
01-WAY-2000 (TrEMBLrel. 13, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                      83.5%; Score 851.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_caps1d.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AF177040; AAF01182.1; -. AF177038; AAF01180.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4b
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               82.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 82.48
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IPR001490;
                                                                                                                                                                                                                                                                                                                                                                   3011 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             P27958; 1HEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID-11103;
```

```
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                               InterPro;
InterPro;
InterPro;
                                                                                                  HSSP; P279
InterPro;
                                                                                                                                                                                                                           InterPro;
InterPro;
                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 10
091RR8
AC 091RR8
DT 01-DEC
DT 01-DEC
DT NS3 PAI
DE NS3 PAI
OC VIVUS
CC HEPAIL
OC HEPAIC
OC NEBICO
OX NCBI
SO THE STANFORM OF THE STANFOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bukh J.;
Subatted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
-1-SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSTE: NOO487; DEXDC: 1.
PROSTE: PS00392: DEXDC: 1.
PROSTE: PS00392: DEX GAD. HDC_YDC; 1.
PROSTE: PS50507: RDRP_POSITIVE; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein: SEQUENCE 3015 AA; 328159 MM; B7D23BC1F190663A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VIruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a second major
and 2a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 12; Length 3015;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Pred. No. 1e-72;
10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-99420396; PubMed-10489358;
Yanagi M., Purcell R.H., Berrson S.U., Bukh J.;
Yahapatlis C virus; an infectious molecular clone of
genotype (2a) and lack of viability of intertypic la
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         83.5%; Score 851.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR00119; Pyridoxal_dec.
InterPro; IPR00119; Pyridoxal_dec.
InterPro; IPR001199; RNA_pol_DS_PS.
InterPro; IPR001099; RNA_pol_DS_PS.
InterPro; IPR001099; RNA_pol_DS_PS.
InterPro; IPR001199; Pyridoxal_dec.
IPfam; PF01540; HCV_capsid: 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 STRGVAKAVDFIPVESLETTMRSP 197
InterPro; IPR001650; Helicase_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Virology 262:250-263(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity 82.4 tes 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     chimeras.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 9
Q9PWU9
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TO DESCRIPTION OF SERVICE OF SERV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
```

```
3 KKGSVVIVGRIN-----LSCDTAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; saRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    83.5%; Score 851.5; DB 12; Length 3015; 82.4%; Pred. No. 1e-72; Ive 10; Mismatches 17; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART: SM00487; DEXDC: 1.

PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.

PROSITE; PS50507; RDRR_POSITIVE; 1.

PROSITE; PS5051; RDRP_VIRAL; 1.

ATP-binding; Cat protein; Envelope protein; Glycoprotein; Hydrolase; Nonstructural protein; Polyprotein;

FNA-directed RNA polymerase; Transferase; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ymerase; Transferase; Transmembrane.
328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181 AA
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 091RR8;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                         | IPR002868; HCV_NS5a.
| IPR00166; HCV_RGRP.
| IPR001650; Helicase_C.
| IPR00129; Pyridoxal_deC.
| IPR007095; RNA_DOL_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01543; HCV_capsid; Pfam; PF01542; HCV_capsid; I. Pfam; PF01560; HCV_capsid; I. Pfam; PF01560; HCV_NS1; I. Pfam; PF0158; HCV_NS2; I. Pfam; PF01006; HCV_NS4; I. Pfam; PF01001; HCV_NS4b; I. Pfam; PF01001; HCV_NS4b; I. Pfam; PF00271; helicase_C; I. Pfam; PF00951; helicase_C; I. Pfam; PF00998; Viral_RARP; I. PF0D0m; PD186062; HCV_NS1; I. PF0D0m; PD186062; HCV_NS1; I.
                                                                                       IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                    IPR002518; H
IPR004109; H
IPR000745; H
IPR001490; H
                                                                                                                                                                               IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE 3015 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
```

~

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4 TAYAQOTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 GSLLSPRPISYLRGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19 TAYAQQTRGEQGCQRTSHTGRDRNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                               Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genedic Diversity and response to IrN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF805238; AAR54563.1;
InterPro: IPR004109; HCV-NS3.
PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from clinical Strains of the Hepatitis C Virus.";
Submitted (APR-201) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369241; AAK54566.1;
Interpro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 181
181 AA; 19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 181
181 AA; 19084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                              O91RR5 PRELIMINARY; PRT; 181 AA.
091RR5.
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 proclease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 847; DB 12;
Pred. No. 7.2e-74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           83.0%;
92.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23, NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Pt. 4V;
                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-Pt. 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SECTENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              091RR2
091RR2;
                                                                  RESULT 12
Q91RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           091RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        64 IASPKGPVIQMYTNVDKDLVGMPAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4 TAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                               4 TAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19 TAYAQQIRGEQGCQKISHIGRDKNQVEGEVQIVSIAIQIFLAISINGVLWIVYHGAGIRT
                                        STRAIN-Pt.1x;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to Irbn of the NS3 Protease Gene from
Submitted Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF362235; AAK54560.1;
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.

RMBL, AF865218, AR854311:
-InterPro; IPR004109; HCV_NS3.

Pfam; PF02907; HCV_NS3; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 849; DB 12; Length 181;
Pred. No. 4.6e-74;
1; Mismatches 11; Indels (
                                                                                                                                                                                                                                                                                                                                     ; DB 12; Length 181;
4.6e-74;
                                                                                                                                                                                                                                                                                                                                                                                            11; Indels
                                                                                                                                                                                                                                                                           181
19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 1
181 181
181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                            Score 849; DB Pred. No. 4.6e-1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                              83.2%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          83.2%;
93.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     166; Conservative
                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 93.3
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NS3 protease (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepacivirus
                                                                                                                                                                                                                                                    NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        139
                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Protease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091RT5
  ACCOCC OST REPRETED BY THE STATE OF THE STAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
```

```
NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                Protease.
NON_TER
NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                 79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                        64 IASPKGPVIQNYTNVDQDLVGWPAPQGARSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78
                                                                                                                                                                                    78
                                                                                                                                                                                                                               4 TAYAQQTRGLLGCIXISLIGRDRNQVEGEVQIVSTAXQTFLATCINGVCWTVYHGAGIRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Pt.1C;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Genetic Diversity and response to IFN of the NS3 Protease Gene from
"Genetic Diversity and response to IFN of the EMBL/GenBank/DDBJ databases.

EMBL, AF369214; AAK54539.1;

ThterPro; IPR004109; HCV.NS3.

Pfam; PF02907; HCV_NS3; I.
                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                         0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.0%; Score 847; DB 12; Length 181; ilarity 92.7%; Pred. No. 7.2e-74; Conservative 2; Mismatches 11; Indels
                           Length 181;
                                                                                                         11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 181
181 AA: 19114 MW; BEIDOB542F014E86 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-DEC-2001 (TTEMBLrel. 19, Created)
01-DEC-2001 (TTEMBLrel. 19, Last sequence update)
01-WAR-2003 (TTEMBLrel. 23, Last annotation update)
                       Score 847; DB 12;
Pred. No. 7.2e-74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AA
                                                                                                         2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      091RR3;
01-DEC-2001 (TrEMBLrel. 19, Ct
01-DEC-2001 (TrEMBLrel. 19, Lz
01-MAR-2003 (TrEMBLrel. 23, Ls
NS3 procease (Frament).
Hepatitis C virus.
                               83.0%;
92.7%;
                                                                          Best Local Similarity 92.7
Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
nes 165; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         19
                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q91RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             091RT9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 14
091RT
AC 091RT
AC 091RT
DT 01-DB
DT 01-DB
DT 01-MS
DT 01-DS
DT 01-MS
DT 01-DS
DT 01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
091RR3
1D Q91RR DT 01-DE
DT 01-DE
DT NO3 P
DE NS3 P
DE NS4 P
DE N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

```
ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19 TAYAQQTRGEQGCQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-Pt.4B;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; AF369240; AAK54565.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                        DB 12; Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                         1 1
181 181
181 AA: 19115 MW; 5D85F88AD7AC1A11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                  82.9%; Score 846; DB 12; 1
92.7%; Pred. No. 9e-74;
tive 2; Mismatches 11;
                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 92.74
Matches 165; Conservative
SEQUENCE FROM N.A.
```

Search completed: August 30, 2003, 19:18:21 Job time : 38.5921 secs

Run on:

```
AR145257 Sequence
AR145255 Sequence
AR145270 Sequence
AR145270 Sequence
AR145275 Sequence
AR145256 Sequence
AR145256 Sequence
AR145251 Sequence
AR145273 Sequence
AR145273 Sequence
AR145273 Sequence
AR145275 Sequence
AR145275 Sequence
AR145279 Sequence
AR145279 Sequence
AR14529 Sequence
AR14529 Sequence
AR14529 Sequence
AR147810 Sequence
AR177810 Sequence
AR177810 Sequence
AR177810 Sequence
BD081910 Hepatitis
AR177810 Sequence
BD081910 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                106434 Sequence 48
109328 Sequence 8
AR118692 Sequence
106440 Sequence 54
109329 Sequence 10
AR118696 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR118696 Sequence
109331 Sequence 15
M32084 Hepatitis C
AR118703 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AF268278 Pestiviru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AF009606 Hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AR145258 Sequence
AR145253 Sequence
AR145266 Sequence
AR145261 Sequence
AR145269 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR110831 Sequence
BD069985 Functiona
AR118686 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                        AR179057 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AR145262 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ALIGNMENTS
                                                                                                                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145272
AR145256
AR145250
AR145251
AR145273
AR145273
AR145273
AR145273
AR145259
AR127810
AR127810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BD081910
AR110828
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR118686
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR118692
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR110831
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BD069985
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R118696
                                                                                                                                                                                                        em_htgo_other:*
                                           em_htg_other:*
                                                                                                                                                                 em_htgo_hum:*
em_htgo_mus:*
em_htg_hum:*
em_htg_inv:*
                                                                            em_htg_pln:*
                                                                                                                em htg mam: *
                                                                                                                                   em_htg_vrt:*
                                                            em_htg_mus:*
                                                                                                                                                                                                                                                                                                                                                                      DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8157
8157
1932
1932
9646
9646
9646
12980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5360
5360
6785
6785
6785
7310
7310
7310
                                                                                                                                                                                                                                                                                                                                                   Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   651
1998
1998
2016
12734
651
651
1998
2016
648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   892.5
888.5
8881.5
8881.5
878.5
878.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.5
877.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  868
867.5
867.5
866.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           866
864
863.5
863.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Command line parameters:
-WODEL-frame+p2n.model -DEV-x1p
-WODEL-frame+p2n.model -DEV-x1p
-WODEL-frame+p2n.model -DEV-x1p
-WOOPEL-frame+p2n.model -DEV-x1p
-DE-GENEDL -OFFT-fastap -SUFFIX-ege -WINMATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-b1ts -START-1 -END--1 -WATRIX-blosum62 -TRANS-human40.cd1 -LIST-45
-DOCALIGN-200 -THR_SCORE-pct -THR_MAX-100 -THR_MIN-0 -ALIGN-15 -MODE-LOCAL
-UNENT-ptc -NORM-ext -HEAPSIXE-500 -THR_MIN-0 -MAXLEN-200000000
-USER-GEO9965594_CGCN.1 1.14686_RTUNAt_29080001.151919_28310 -NCPU-6 -ICPU-3
-NO_MARP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONCLOG
-DEV_INMOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
                                                                                                      August 30, 2003, 19:18:33 ; Search time 2560.57 Seconds (without alignments) 3147.423 Million cell updates/sec
                                                                                                                                                                                                      1020
1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP 197
   GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                            OM protein – nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                    2888711 segs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                                                                                                      of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                    Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                        Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                    US-09-965-594-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   90.ba:

90.htg:

90.ni:

90.on:

90.pat:

90.pat:

90.pl:

90.pl:

90.pr:

90.pr:

90.pr:

90.pr:

90.pr:

90.pr:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_pat:*
em_ph:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 em_pl:*
em_ro:*
em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_fun:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  em_in:
em_mu:
em_om:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   em_hum: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              *: vo_ma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GenEmbl: *
                                                                                                                                                                                                                                                           BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Title:
Perfect score:
                                                                                                                                                                                                                                                           Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                        Total number
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Database :
                                                                                                                                                                                                                       Sequence:
                                                                                                                                                                                                                                                                                                                                                                      Searched:
```

```
AR145264
AR145264.1
                 .
Unknown.
                               Unknown
                                                                                                                          Scores:
                                                                                                                                                                                                                                                                                                                                                102
                                                                                                                                                                                                                                                                                                                                                                            122
                                                                                                                                                                                                                                                                                                                                                                                         424
                                                                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                                     162
                                                                                                                                                                                                                                                                                                                   82
                                                                                        source
                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                   ...
No.
ACCESSION
VERSION
                              ORGANISM
                                                   AUTHORS
TITLE
                                                                                                                          Alignment
                                             REFERENCE
                                                                        JOURNAL
               KEYWORDS
                                                                                FEATURES
                       SOURCE
                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                             줨
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                  ð
                                                                                                                                                                                                               g
                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                   å
                                                                                                                                                                                                                                                                                                                                                                                                                     20-APR-2002
                                                                                                                                                                                                                                                                                                                                592
                                                                                                                                                                                                                                                                                                                   652
                                                                                                                                                                                                                                                                                                                                                                                          ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                             41
                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                              AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                          AlaginginthrArgGlyGluGinGlyCysGluLysThrSerHisThrGlyArgAspLys
                                                          Unclassified.

1 (bases 1 to 12734)

Hong,Z., Lai,V.C.H. and Lau,J.Y.N.

Hepatitis 0 virus protease-dependent chimeric pestivirus

Patent: US 6326137-A 1 04-DEC-2001;

Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
         DNA
       AR179057 12734 bp
Sequence 1 from patent US 6326137.
                                                                                                                                                                                                 (1-12734)
                                                                                                            /organism="unknown"
1 2604 c 3295 g
                                                                                                                                                                                                  US-09-965-594-20 (1-197) x AR179057
                              AR179057.1 GI:20220612
                                                                                                                                               1.5e-65
892.50
92.82%
90.77%
87.50%
                                                                                                                                                              Percent Similarity:
Best Local Similarity:
                                                                                                                     ď
                                             Unknown.
                                                    Unknown
                                                                                                                                          Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                            713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  893
                                                                                                                                                                                                                                             22
                                                                                                                                                                                                                                                                                       533
                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                    162
                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                   593
                                                                                                                                                                                                                                                                                                                                                553
                                                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                            Query Match:
DB:
                                                                                                       source
AR179057
LOCUS
DEFINITION
                      ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                   BASE COUNT
ORIGIN
                                                                        AUTHORS
TITLE
JOURNAL
                                                                  REFERENCE
                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                            ŏ
```

08-AUG-2001

PAT

linear

AR145264 1998 bp | Sequence 105 from patent US 6211338.

AR145264 LOCUS DEFINITION

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlabroginglySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                 22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLy8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        544 reaggeracerergescarcrecegecrecerargeaccegegegerangec
                                                                                                                                                                                                                                                                                                                                                                        GlySerValVallleValGlyArgIleAsnLeuSerGlyAsp...--ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,
Single-chain recombinant complexes of hepatitis C
protease and NS4A cofactor peptide
Patent: US 6211338-A 105 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1998 bp DNA Sequence 109 from patent US 6211338. AR145268 1 GI:15107135
                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                             US-09-965-594-20 (1-197) x AR145264 (1-1998)
                                                                                                                                                                      /organism="unknown"
595 c 569 g
                                                                                                                                              Location/Qualifiers
GI:15107131
                                                                                                                                                                                                                                        6.96e-66
885.50
93.378
85.718
86.818
                                                                                                                                                          1998
                                                          Unclassified.
                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 3
AR145268
LOCUS
```

m

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                        243
                                                                                                                                                                                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    603
                                                                                                                                                                                                                                    4
                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                          244 TGCGTCAACGGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                           22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                             NS3
                     (bases 1 to 1998)

Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NSAA cofactor peptide
Patent: US 621138-A 104 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    651 bp
S 6211338.
                                                                                                      424
                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                 US-09-965-594-20 (1-197) x AR145263 (1-1998)
                                                                                        /organism~"unknown"
596 c 568 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145254 65
Sequence 95 from patent US
AR145254
                                                                                                                                            1.52e-65
881.50
92.86%
85.71%
86.42%
                                                                                    .1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown.
Unclassified
            Unclassified
                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR145254.1
                                                                                                     410 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown
                                                                                                                                    Scores:
                                                                                                                                                                                     Query Match:
                                                                                    source
                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                      BASE COUNT
   ORGANISM
                                AUTHORS
TITLE
                                                                                                                                   Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
AR145254
                                                                        FEATURES
                                                                                                                ORIGIN
                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                          á
                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                             合
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                423
                                                                                                                                                                                                                                                                                                                                                                                                 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                    ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                41
                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                      22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                             SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                        1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus protease and NSAA cofactor peptide
Patent: US 6211388-A 109 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                               1998
167
16
10
3
                                                                                                                                                                        Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1998 bp I
from patent US 6211338.
                                                                                                                                                                Length:
                                                                                                                                                                                                                                   US-09-965-594-20 (1-197) x AR145268 (1-1998)
                                                                                                             /organism="unknown"
595 c 569 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 104 from paten:
AR145263
AR145263.1 GI:15107130
                                                                                                                                                             1.25e-65
882.50
93.37%
85.20%
                                                                                                      1. .1998
                                Unclassified
                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                        ಡ
       .
Unknown .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown
                     Unknown
                                                                                                                                                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                      42
                                                                                                                                                                                                                                                                                                                                                                             62
                                                                                                                                                                                                                                                                                                                                                                                                                                                            102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182
                                                                                                                                                                                                                                                                                                                                                                                                                    82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 4
AR145263
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                       Query Match:
                                                                                                      source
                                                                                                                                                               Pred. No.:
                      ORGANISM
                                                                                                                        BASE COUNT
                                        REFERENCE
AUTHORS
TITLE
                                                                                JOURNAL
FEATURES
KEYWORDS
                                                                                                                                  ORIGIN
```

q õ 셤 ŏ a ò g ò 셤 ò a ò g õ

ò

g ö 셤 ò a

```
Unclassified
                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown
                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                         62
                                                                                                                                                                                                                                                                                                                                                83
                                                                                                                                                                                                                                                                                                                                                                  304
                                                                                                                                                                                                                                                                                                                                                                                     107
                                                                                                                                                                                                                                                                                                                                                                                                       364
                                                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                             124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         544
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
ACCESSION
                                          source
                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
TITLE
                                                           BASE COUNT
                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
AR145262
                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
   TITLE
                                                                     ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                         g
                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                                                                                                              141
                                                                                                                                                                                                                                                                                                                                                                                                                                                        GCGCCCCCCGGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTACTTG 423
                                                                                                                                                                                                                            64 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                         GlyServalValIleValGlyArg1leAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                                                                                                AlaGlnGlnThrArgGlyGluGlnGlyCy8GlnLy8ThrSerHisThrGlyArgAspLy8 41
                                                                                                                                                                                                                                                                                                                            SerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                                                                               TGCGTCAACGGCGTGTTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                       AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
|||||||
                                                                                                                                                                                                                                                                                                                                                                                                                                           ValThrArgHisAlaAspValIleProValArgArgArgClyAspSerArgClySerLeu
                                                                                                                                                                                                                                                                                     AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4 cofactor peptide
Patent: US 6211338-A 95 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              linear
                                                                                                                    651
167
15
3
                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 108 from patent US 6211338.
AR145267
AR145267.1 GI:15107134
                                                                              144
                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1998 bp
                                                                                                                                                                                       US-09-965-594-20 (1-197) x AR145254 (1-651)
                                                 Location/Qualifiers
                                                                    /organism="unknown"
187 c 200 g
                                                                                                                   7.52e-66
878.50
93.33%
85.64%
86.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unknown.
Unclassified.
                                                                                                                                     Percent Similarity:
Best Local Similarity:
Query Match:
                                                                               Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AR145267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown.
                                                                                                           Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                          62
                                                                                                                                                                                                                                                                                                                                                244
                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                    304
                                                                                                                                                                                                                                                                                                                                                                                                       102
                                                                                                                                                                                                                                                                                                                                                                                                                         364
                                                                                                                                                                                                                                                                                                                                                                                                                                                               424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    484
                                                                                                                                                                                                                                                22
                                                                                                                                                                                                                                                                                     42
                                                                                                                                                                                                                                                                                                                                                                                                                                             122
                                                            source
                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 6
AR145267
  REFERENCE
            AUTHORS
                                        JOURNAL
                                                  FEATURES
                     TITLE
                                                                                                                              Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                         ö
                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                 á
                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                             οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  š
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 GGTTCTGTTGTTATTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                         CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                          41
                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                          GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                        22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                      42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Single-chain recombinant complexes of hepatitis C protease and NS4A cofactor peptide
Patent: US 6211318-A 108 03 APR-2001;
Location/Qualifiers
1. 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                          1998
167
15
11
                                                                                                                                                                             Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 103 from patent US 6211338. AR145262
                                                                                               424
                                                                                                                                                                                                                              Indels:
                                                                                                                                                              Length:
                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                           US-09-965-594-20 (1-197) x AR145267 (1-1998)
                                                                               /organism="unknown"
596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AR145262.1 GI:15107129
                                                                                                                                                         2.73e-65
878.50
92.86%
85.20%
86.13%
```

S

```
Percent Similarity:
Best Local Similarity:
                   120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                                 Aliqnment Scores:
                                                                                                                                                                                                                                                                                                     82
                                                                                                Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                  BASE COUNT
ORIGIN
                                                          ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PEATURES
                                                           Pred
                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                              ò
                                                                                                                                                                ద
                                                                                                                                                                                    ö
                                                                                                                                                                                                   셤
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                            LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            243
                                                                                                                                                                                                                                                                                                                                                                                                  ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG 603
                                                                                                                                                                                                                                                                                                                      ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                            AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                      AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                                           SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                   244 TGCGTCAACGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                  GlySerValValIleValGlyArg1leAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 (bases 1 to 651)
Malcolm.B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211138-A 99 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
6211338-A 103 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AR145258 651 bp
Sequence 99 from patent US 6211338.
AR145258
                                      424
                                                                                                                            Gaps:
                                                                                                                                             (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                          /organism="unknown"
1 596 c 568 a
          Location/Qualifiers
                                                                                                                                              US-09-965-594-20 (1-197) x AR145262
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145258.1 GI:15107125
                                                                            3.33e-65
877.50
92.86%
85.20%
86.03%
                  .1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unclassified.
                                                                                              Percent Similarity:
Best Local Similarity:
Patent: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unknown
                                                                   Alignment Scores:
                                                                                                                                                                                                                         124
                                                                                                                                                                                                                                            43
                                                                                                                                                                                                                                                                                62
                                                                                                                                                                                                                                                                                                                                                           102
                                                                                                                                                                                                                                                                                                                                                                              364
                                                                                                                                                                                                                                                                                                                                                                                                  122
                                                                                                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            162
                                                                                                                                                                                                                                                                                                                                                                                                                    124
                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
AR145258
LOCUS
DEFINITION
ACCESSION
KEYWORDS
SOURCE
                     source
                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            JOURNAL
FEATURES
           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                 ò
                                                                                                                                                                                   염
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                  141
                                                                                                                                                                                                                                                                                                                                                                                                               483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        161
                                                                                                                                                       ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                   22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                           GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Unclassified.

1 (bases 1 to 651)

Maloolin,B.A., Tareml,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide
Patent: US 6211338-A 94 03.APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            651
166
16
10
1
                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145253 651 bp C
Sequence 94 from patent US 6211338.
AR145253 GI:15107120
           144
                                                                                                                         (1-651)
/organism-"unknown"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /organism-"unknown'
            σ
         200
                                                                                                                         US-09-965-594-20 (1-197) x AR145258
                                                1.35e-65
875.50
93.33%
85.13%
85.83%
         187 c
```

119 a 188 c 199 g 145 t	Alignment Scores:	Screet Similarity: Secret Similarity: Percent Similarity: Best Local Similarity: Query Match: DB: US-09-965-594-20 (1-197)	Qy 5 Glyservalvalile	21 Db 64	STAGTATCACGGCCTAC 123 Qy 2.2 ALBGINGINTERFY :::	183 Oy 42	61 Db 184 243 Qy 62	81 Db 244 TGCGTCAAC 303 Qy 82 PTOLYSGIY		Actrodrogologias and the property of the programme of the property of the prop	423 Qy 122		483 Oy 142	161 Db 484 CT 543 Qy 162 Al	:::	603 Oy 182	196 648	LOCUS Linear PAT 08-AUG-2001 CERSION AR145261 ACCESSION AR145261 ACCESSION AR145261 KEYWORDS UNKNOWN. ORGANISM UDKNOWN.	REFERENCE 1 (bases 1 to 1 t	Source Source
	c 199 g 145		(1-197) x AR145253 (1-651)	GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp	GTTATTGTTGGTAGAATTATTTTATCTGGTAGTGG ThrarqG1yG1uG1nG1yCysG1nLysThrSerHi	:::	AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 	SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer:::	ProValThrGlnMetTyrThrAsnValAspLysAs	CCAATCACCCAGATGTACATTATGTGGACCAGA GlySerArgSerLeuthrProCysThrCysGlySe	GGGGGGGTTCCTTGACACCATGCACCTGTGGCAG	HisAlaAspvalIleProvalArgArgArgGlyAs	ACATGCTGACGTCATTCCGGTGCGCCGGCGGGCGA	LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 	sAlaValGlyIlePheArgAlaAlaValSerThrAr	TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGGTTGCGAAGGCG	ValasppheileProValGluSerLeuGluThrThrMetArgSer 	1998 bp DNA nt US 6211338. 3	Unknown. Unknown. Unknown. Unklossified. 1 (bases 1 to 1998) Malcolm.B.A., Taremi.S.Shane., Weber, P.C. and Single-chain recombinant complexes of hepatit. Protease and NS4A cofactor peptide	S 6211338-A 107 03-APR-2001; Location/Qualifiers 11998

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                            AASPVAllleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                      yGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
||||||||||||||
AGAGGTTCAGGTGGTTTCCACCGCAACAATCCTTCCTGGGGACC 243
                                                                                                                                                                                                                   eValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               aremi, S. Shane., Weber, P. C. and Yao, N. combinant complexes of hepatitis C virus NS3 4A cofactor peptide 338-A 102 03-APR-2001; on/Qualifiers 98 15m="unknown" 7 c 567 g 425 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      roValGluSerLeuGluThrThrMetArgSerPro 197
|||||||||||||
|CGTAGAGTCCATGGAACTACTATGCGGTCTCCG 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
1998
166
16
11
 Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1998 bp DNA om patent US 6211338.
                                                                        R145266 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15107128
```

Percent Similarity: 92.35% Conservative: 14 Best Local Similarity: 85.20% Mismatches: 12 Query Match: 85.64% Indels: 3 DB: 6 Gaps: 1 US-09-965-594-20 (1-197) x AR145269 (1-2016)	Qy 5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAlaTyr 21	Qy 42 AsnGlnValGluUalgGluValGlnIleValSerThrAlaThrGlnThrPheLeualaThr 61 Db 202 AACCAGGTCGAGGAGAGGTTCAGGTGGTTTCCACCGCAACATCCTTCCT	Qy 82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLeuValGlyTrpGln 101	0y 122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141	Qy 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181 Db 562 TCGGGGCACGTGGGCATCTTCCGGGCTGCCGTATGCACCGGGGGTTGCGAAGGCG 621 Qy 182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197 Db 622 GTGGACTTTGTGCCCGTAGGAACTACTATGCGGTCTCCG 669 RESULT 13	AF268278 AF268278 LOCUS DEFINITION PESTIVILUS type 1, complete genome. ACCESSION AF268278 AF268278.1 GI:9049956 VERSTON KEYWORDS SOURCE ORGANISM VINUESS: SRRNA positive-strand viruses, no DNA stage; Flavivilidae; PESTIVILUS PESTIVILUS TITLE ORGANISM AUTHORS DOMIS, R.O., Hong, W., Skelton, A., Ingravallo, P., Vassilev, V., Domis, R.O., Hong, Z. and Lau, J. Y. TITLE Generation and characterization of a hepatitis C virus NS3 DOUGHALL J. VIIOL, 74 (14), 6339-6347 (2000)	g.
Pred. No.: 7.27e-65 Length: 1998 Score: 873.50 Matches: 167 Percent Similarity: 92.35% Conservative: 14 Best Local Similarity: 85.20% Mismatches: 12 Query Match: 6 Gaps: 1	a − o • •	Db 124 TCCCAACAGACGGGCCTACTTGGTTCATCTTACTAGGGCGGACAAG 183 Qy 42 AsnGlNVA1GLUGLYGLUVA1GLNIEVA1SATTACATAGCTTACAGGCCGGGACAAG 183 Qy 42 AsnGlNVA1GLUGLYGLUVA1GLNIEVA1SATTACATATATATATATATATATATATATATATATATAT	Db 244 TGCGTCAACGGCGTGTGTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303 Qy 82 ProLysGlyProvalThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101	Db 364 GCGCCCCCGGGGGGGCGTTGTTGTTTTTTTTTTTTTTT	Db 484 CTCTCCCCAGGCCTGTCTCTACTTGAAGGCTCTTCGGGTGGTCCTCTCCCCT 543 Qy 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181 Cy 11	RESULT 12 AR145269 LOCUS LOCUS DEFINITION Sequence 110 from patent US 6211338. ACCESSION AR145269 VERSION AR145269 VERSION AR145269.1 GI:15107136 KEYWORDS SOURCE ONGANISM GINCHOWN. GINCHOWN. GINCHOWN. TITLE SINGle-Chain recombinant complexes of hepatitis C virus NS3 FEATURES TOTAL PATENT. FEATURES LOCATION ON A 10 03 4211338 A 110 03 APR-2001; FEATURES LOCATION ON 11 fiers LOCATION ON 11 fiers	NT 412 a t Scores:

œ

organism-"Pestivirus

5'UTR CDS

```
TEGIORMECINOVIRGING MERICING PRILATION TO REPROTECT TO THE STORMEDISCHEDISCH THE STORMEDISCH T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LVFVPTRIMAVEVAKKLKAKCYNSCYYSGEDPANLRYVTSGSPYVTYATNI LESGVY
LPDLDTVIDTGLKCERKVRYSSKIPFIYTGLKRANVYTSGSPYVTYATNI LESGVY
LPDLDTVIDTGLKCERKVRYSSKIPFIYTGLKRANVYTSGSPYVTYATNI LESGVYY
LLGSEDPAVKNIHARTDHFEPIOLAYNSFERNYTWEDAGDRORYGRYRSFLN
ARKLGEDVPVYIYATEDEDLAVDLLGLDWPDPGNQOVVETGKALKOVTGLSSAENALL
ARKLGEDVPVYIYATEDEDLAVDLLGLDWPDPGNQOVVETGKALKOVTGLSSAENALL
GDVEVYYQOALSKHWPMITDIYITEDQRLEDTTHLQYABNAIKTDGTETELKELAS
GDVEKINGAISDYAAGGLEVKSQARKIKTAPLFKENAEAARGYVOKFIDSLIENKEE
IIRYGLWGTHTALY KSIAARLGHETAFATLVLKWLAFGGESVSDHYKQAAVDLLVVYYY
MNKFSFPGDSETQOEGRRFVASLFISALATYKTWNYHNLSKVVEPALAXLLPYATSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AYTEVNEDVGTIK PYLGERVIPDPVVDINLOPEVQVDTSEVGITIIGRETLMTTGYTP
VLERVEPDASDNONSVRIGLDEGNYPGPGIOTHTLTEEIHNRDARPFIMILGSRNSIS
NRAKTARNINLYTGNDPREIRDLMAAGRMLVVALRDVDPELSEWVDFKGTFLDREALE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LKMFTPTRLESVVILSTTIYKTYLSIRKGKSDGLLGTGISAAMEILSQNPVSVGISVM
LGVGAIAAHNAIESSEQRRTLLMRVPVKNFLDQAATDELVKENPERIIMALFEAVQTI
GNPLRLIYHLYGVYYKGWEAKELSERTAGRNLFTLIMFEAFELLGMDSQGKIRNLSGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ALSLGQPKPKQVTKEAVRNLIEQKKDVEIPNWFASDDPVFLEVALKNDKYYLVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AGIRLEKLPIVRAQTDTKTFHEAIRDKIDKSENRÖNPELHNKILLEIFHTIAOPTLKHT
YGEYTWEQLEAGINRGAAGFLEKKNIGEVLOSEKHLVEOLVROLKAGRKIKYTETAI
PKREKRDVSDDWAGGLVVEKRPVIOYPEAKTRAITKVHYNWYRQPVYIEYEGY
TPLFNIFDKVKKEWDSFNEPVAVSFDTRAMDTQVTSKDLQLIGEIQKTYYKKEWHKFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMKAFKNVGGKLTKVEESGPFLCRNRPGRGPVNYRVTKYYDDNLREIKPVAKLEGQVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HYYRGVTAKIDYSKGKMLLATDKWEVEHGVITRLAKRYTGVGFWGAYLGDEPNHRALV
ERDCATITRNTVQFLKWKKGCAFTYDLJISNLTRLIELVHRNNLEEKEIPTATVTTWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             VKDQAKALGATDQTRI IKEVGSRTYAMKLSSWFLQASNKOMSLTPLFEELLLRCPPAT
KSNKGHWASAYQLAQGNWEPLGCGVHLGTIPARRVKIHPYEAYLKLKDFIEEEEKKPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VKDTVIREHNKWILKKIRFQGNLNTKKMLNPGKLSEQLDREGRKRNIYNHQIGTIMSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SYSDTKEEGATKKKTOKPDRLERGKMKIVPKESEKDSKTKPPDATIVVEGVKYQVRKK
GRTKSKNTQDGLYHNKNKPQESRKRLEKALLAWAIIAIVLFQVTMGENITQWNLQDNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YILDLIYGLHKQINRGLKKMVLGWAPAPFSCDWTPSDERIRLPTDNYLRVETRCPCGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /translation-"melnTheGSGSVVIVGRIVLSGSGSITaCAQQTRGLLGCKITSL
TGRDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGAGTRTIASPKGPVIQMYTNVD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QDLVGWPAPQGSRSLTPCTCGSSDLXLVTRHANVIPVRRRGDSRGSLLSPRPISYLKG
SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TEGIORAMFORGVINSLHGIMPEKICTGVPSHLATDIELKTIHGMMDASEKTNYTCCR
                                                                                                                                                                                                                                                                                            /product-"polyprotein"
/protein_id-"AAF82566.1"
/db_xref-"G1:9049957"
/mol_type="genomic RN;
/db_xref="taxon:11099
                                                                                                                                                                                                                                            /codon_start-1
                                                                                                                               .385
```

```
DIIDHMTEVPVITADGEVIIRNGQRGSGQPDTSAGNSMLNVLTMMIAFCESTGVPYK
SFNRVARIHVGGDDGFLITEKGLGLKFANKGWOILHEAGKPOKITEGERMKVAYRFED
FECSHTPPVPVMSDNTSSHMAGNDFVILSKMATRLDSGERGTTAFERAVAFSFLL
MYSKNELVRRICGLLVLSQQPSTDPSKRATYYKGDPIGAYKDVIGSRLSELKHTGFER
LANLNLSLSTLGIWTKHTSKRIIQDCVAIGKEEGNWLNAADRLISSKTCHLYIPDKGF
TLQGRHYEGGLLKTETNPVMGVGTERYKLGPIVNLLLRRLKILLMTAVGVSS*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                           592
                                                                                                                                                                                                                                                                                                                                                                                                                    772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                   AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                       41
                                                                                                                                                                                                                                                                                                                 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerIeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GCTCCTCAAGGTTCCCGCTCATTGACACCCTGCACCTGCGGCTCCTCGGACCTTACCTG
                                                                                                                                                                                                                                                                     22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 (bases 1 to 651)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis G virus NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                           Н
                                                                                                                                  Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA
                                                                             Ų
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 98 from patent US 6211338.
AR145257
AR145257.1 GI:15107124
                                                                           2802
                                                                                                                        Length:
                                                                                                                                                                                                    (1-12734)
                                                                                                                                                                              Gaps:
                                                                              D.
                                                                             3293
                                                                                                                                                                                                     US-09-965-594-20 (1-197) x AF268278
                                                                  12509. .12734
2608 c 33
                                                                                                                       7.45e-64
872.50
91.79%
89.23%
85.54%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unclassified.
                                                                                                                                        Percent Similarity:
Best Local Similarity:
Query Match:
                                                                              ಹ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown.
                                                                             4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown
                                                                                                              Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182
                                                                                                                                                                                                                                                113
                                                                                                                                                                                                                                                                                                                42
                                                                                                                                                                                                                                                                                                                                       533
                                                                                                                                                                                                                                                                                                                                                                                                                                                   102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          893
                                                                                                                                                                                                                           S
                                                                                                                                                                                                                                                                                           473
                                                                                                                                                                                                                                                                                                                                                                                                        82
                                                                                                                                                                                                                                                                                                                                                                                                                             653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      953
                                                                 3'UTR
BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 14
AR145257
                                                                                                              Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                       ORIGIN
                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ם
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                            8
                                                                                                                                                                                                                                                                                            윱
                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                             ò
```

Location/Qualifiers

```
Scores
                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                42
                                                                                                                                                                                                                                                                      102
                                                                                                                                                                                                                                                       104
                                                                                                                                                                                                                                                                                                                  424
        source
                      BASE COUNT
ORIGIN
                                                     No.
                                              Alignment
 FEATURES
                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                             á
                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                ογ
                                                                                                                      ö
                                                                                                                                   g
                                                                                                                                                              a
                                                                                                                                                                               ò
                                                                                                                                                                                            g
                                                                                                                                                                                                         οχ
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                         ò
                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAT 08-AUG-2001
                                                                                                                                                                                                                                         303
                                                                                                                                                                                                                                                                      363
                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                 483
                                                                                                                                                                                                                                                                                                                                               161
                                                                                                                                                                                                                                                                                                                                                             543
                                                                                                                                                                                                                                                                                                                                                                          243
                                                                                                                                                                                                                                                        101
                                                                                                                                            41
                                                                                                                                                                                              61
                                                                                                                                                                                                                          81
                                                                                                                                                                                             ValThrargfisalaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                              -----ThrAlaTyr
                                                                                                                                                                SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                 GTCACGAGACATGCTGACGTCATTCCGGTGCGCCGGCGGGGGGCGACAGTAGGGGGGAGCCTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 (bases 1 to 651)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4 cofactor peptide
Patent: US 6211338-A 93 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                         182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                       648
                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                651
166
15
11
11
                                                                                                                                     5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-
                                                                  Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA
protease and NS4R cofactor peptide
Patent: US 6211318-A 98 03-APR-2001;
Location/Qualifiers
1. 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 93 from patent US 6211338. AR145252 AR145252.1 GI:15107119
                                                                                                       Gaps:
                                                                                                                      US-09-965-594-20 (1-197) x AR145257 (1-651)
                            /organism-"unknown"
188 c 199 a
                                                                  2.95e-65
871.50
92.82%
85.13%
85.44%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unclassified
                                                                                         Similarity:
                                      Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unknown
                                                                                  Percent Similarity:
                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                 424
                                                                                                                                                                                                                                                                                                                                                                                                                       604
                                                                                                                                                                                                                           62
                                                                                                                                                                                                                                                                                                                  122
                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 15
AR145252
LOCUS
DEFINITION
ACCESSION
VERSION
                                                                                                Query Match:
DB:
                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
                                                                    Pred. No.:
                                      BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             JOURNAL
         JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TITLE
                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                ò
                                                                                                                                                                               a
                                                                                                                                                                                                          g
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                     ò
                                                                                                                                                                                              ö
```

```
243
                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                    141
                                                                                                                                                                                                                                                                                                                                                                                       483
                                                                                                                                                           64 GGTTCTGTTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                    19
                                                                                                                                                                                                                                                       122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                               GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                    AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                 AlabroglnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   651
166
15
11
                                                                        Conservative:
Mismatches:
Indels:
                                                     Length:
Matches:
                  145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search completed: August 31, 2003, 00:46:32 Job time : 2569.57 secs
                                                                                                                       US-09-965-594-20 (1-197) x AR145252 (1-651)
1. .651
/organism="unknown"
                   5
                                                    3.59e-65
870.50
92.82%
85.13%
                 188 c
                                                                                 Best Local Similarity:
Query Match:
                   Ф
                                                                        Percent Similarity:
Best Local Similarit
```

```
AL IGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAV59361
ABK87285
                                                                                                                                                                                                                                                                                                                                                             AAX80348
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX23259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAX23258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABK87286
                                                                                                                                                                                                                                                                                                                                                  AAX80360
                                                                                                                                                                                                                                                                                                                                                                            AAX80343
                                                                                                                                                                                                                                                                                                                                                                                                                  AAX80365
                                                                                                                                                                                                                                                                                                                                                                                                                                            AAX80347
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAX80346
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAN90336
                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX80351
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAX80362
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Location/Qualifiers
                                                       ü
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ВР
                                                                                                                                  AAA73333 standard; DNA; 594
                                                       DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9646
12980
12980
16622
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              5360
6905
7310
7310
                                                                                                                                                                                                                                                                                                           651
1998
1998
2013
651
                                                      Length
                                                                                                                                                                                                                                         651
1998
1998
                                                                                                                                                                                                                                                                                                                                                                           651
1998
2016
648
648
650
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
Synthetic.
                             &
Query
Match 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              888888.58
85888.58
8588.58
8588.58
854.58
864.58
864.58
                                                                                                                                                                                                 885.5
882.5
881.5
878.5
878.5
877.5
877.5
                                                                                                                                                                                                                                                                                                           874.5
874.5
873.5
873.5
871.5
870.5
870.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                      867.5
866.5
864
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             863.5
863.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73333;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   858.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Key
                                                                                                                                                                                                  Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              August 30, 2003, 19:13:57; Search time 182.939 Seconds (without alignments) 2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N. Geneseq_19Jun03:*

| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA20019.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA20019.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA20019.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA20019.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA20019.DAT:*
| SIDSI/gcgdata/geneseq/geneseqn-embl/NA20018.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,
                                                                                                                              US-09-965-594-20
1020
1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /SIDS1/gcgdata/geneseq/genesegn-emb1/NA2003.DAT:*
                                                                                                                                                                                                                                                                                              5105512
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /SIDS1/gcgdata/geneseq/genesegn-emb
                                                    nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                            hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                  2552756 seqs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                 0.5
7.0
7.0
                                                                                                                                                                                                Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                      Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                    BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Total number of
                                                                                                                                  Title:
Perfect score:
                                                                                                                                                                                    Scoring table:
                                                      OM protein -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database :
                                                                                                                                                                                                                                                                  Searched:
                                                                                                                                                             Sequence:
                                                                                on:
```

HCV NS4A-NS3 COMPL HCV NS4A-NS3 Hepatitis C virus Chimetic BVDV/HCV HCV NS4A-NS3 compl is derived by analysis of the total score distribution SUMMARIES AAX80355 AAX80359 AAX80354 AAX80349 ABX15706 AAX80345 AAX80358 AAA73333 ABA95615

Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds. Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.

Combined open read Hepatitis C virus Combined open read Combined open read

Composite hepatiti

Nucleotide sequenc

Hepatitis C virus Hepatitis C virus Hepatitis C virus

CDNA encoding hepa

~

```
GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the POlyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21 TyralaglnGlnThrargGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp 40
                                                                                                                                                                                                                                                                                               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                               /product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 594 BP; 104 A; 191 C; 152 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           594
1197
0
0
0
                                                                                                                                                                                                                     Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-20 (1-197) x AAA73333 (1-594)
                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                           Claim 26; F1g 16; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.17e-87
1020.00
100.00%
100.00%
                                                                                                                          06-JAN-2000; 2000WO-US00345
                                                                                                                                                         990S-0115271
                   Q
              /*tag-
                                                                                                                                                                                                                                                  WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                              WO200040707-A1
                                                                                                                                                                                                                     Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                            13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
```

```
140
                                          420
                                                               160
                                                                                     480
                                                                                                       ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                           481 CCGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTTCCACCCGTGGTGTTGCTAAA 540
301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                               LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                           181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.
                                                                                                                                                                                                                                                                                                                                                                                   /*tag= a
/product= "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y,
                                                                                                                                                                                                                                                                                                                                                              Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 26; Fig 17; 66pp; English.
                                                                                                                                                                                                               BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M, Weinheimer S,
                                                                                                                                                                                                              AAA73334 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0115271
                                                                                                                                                                                                                                                     19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2000-465976/40.
P-PSDB; AAB15225.
                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                         Synthetic
                                                              141
                                                                                                                                                                                                                                   AAA73334;
                    121
                                         361
                                                                                  421
                                                                                                        161
                                                                                                                                                                                                     AAA7333
                                                                                                                                                                                          RESULT
                                                                                                                           a
                                                                                                                                                                     g
                                                                                                                                                                                                                a
                                                                                  a
                                                                                                        ò
                                                                                                                                                ð
                    ò
                                                              ò
```

The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enizymes. These proteins are both wesential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.

180

80

us-09-965-594-20.rng

```
WO200040707-A1
                                                                                                                                                                                                                                                                                                 21
                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                        41
                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                241
                                                                                                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                            101
                                                                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                               Query Match:
DB:
                                                                                                                                                                                                                          ..
9
                                                                                                                                                                                                                                Score:
ò
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                               420
                                                                                                                                                                                                                                               480
                                                                                                                                                                                                                                                         240
                                                                                                                                                                                            180
                                                                                                                                                                                       GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                             LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                    LeuleuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                    9
                                                                                                                                          80
                                                                        50
                                                                                  9
                                                                                              40
                                                                                                    LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                         MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                             TyralaginginThrargglygluginglyCysglnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                              protease; viral replication; chronic liver disease;
liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5
                                                                                                                                                                                                                                                                                                                                                                                                                  /product= "NS4A-NS3 fusion protein #5"
     other;
                     594
196
0
0
0
                    Length:
Matches:
Conservative:
Mismatches:
Indels:
    0
    Ξ;
    G; 146
                                                 Gaps:
                                                            US-09-965-594-20 (1-197) x AAA73334 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
1..594
/*tag= a
    151
                                                                                                                                                                                                                                                                                                                  ВР
    BP; 105 A; 192 C;
                    6.24e-86
1010.00
99.49%
99.02%
                                                                                                                                                                                                                                                                                                                  DNA; 594
                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                  standard;
                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
Synthetic.
                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                     liver failure;
    Sequence 594
                                                                                                                                                                                                                                                                                                                                        19-DEC-2000
               Alignment Scores:
                                                                                             21
                                                                                                                   41
                                                                                                                                                                                      101
                                                                                                                                                                                                 301
                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                               181
                                                                                                                                                                                                            121
                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                 AAA73332
                                                                                                                                                                                                                                                                                                                            AAA73332,
                                     Best Local S
Query Match:
                     Pred. No.:
                                                                                                                                                                                                                                                                                                      RESULT 3
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                 ò
                                                                                 a
                                                                                             ð
                                                                                                       셤
                                                                                                                ò
                                                                                                                             a
                                                                                                                                                                         Q
                                                                                                                                                                                    õ
                                                                                                                                                                                               셤
                                                                                                                                                                                                            ö
                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                            õ
                                                                                                                                          à
                                                                                                                                                   a
                                                                                                                                                               č
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC 360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                40
                                                                                                                                                                                                                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TyralaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MetLysLysLysGlyServalvalIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;
                                                                                                                                                  Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-20 (1-197) x AAA73332 (1-594)
                                                                                                                                                     χ,
                                                                                                                                                  Zhang
                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                   Claim 26; Fig 15; 66pp; English.
                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1.84e-85
1005.00
99.49%
98.48%
98.53%
                                                                                                                                                  Wittekind M, Weinheimer S,
06-JAN-2000; 2000WO-US00345
                                                 990S-0115271
                                                                                                                                                                                               WPI; 2000-465976/40.
P-PSDB; AAB15223.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
```

셤

ద ò

ò 8 ò

```
4.72e-84
990.00
97.97%
96.95%
97.06%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1..594
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA73330;
                                                              Query Match:
             ..
⊗
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Key
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
             Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                              õ
                                                                                                                                                     q
                                                                                                                                                                                 à
                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the Polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. The protein produced from this
                                      160
                                                              480
                                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                      LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyBroLeuLeuCys
            CTGGTTACCCGTCACGCTGACGTTATCCCGGTTCGTCGTCGTGGTGACTCCCGTGGTTCC
                                                              crecreteccescercearcrecratereaaagerrecressiscercegerecretere
Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                         /product- "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wittekind M. Weinheimer S. Zhang Y.
                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                       BP.
                                                                                                                                                                                                                       AAA73331 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0115271
                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-465976/40.
P-PSDB; AAB15222.
                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                         19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                              421
                                                                                                                                                                                                                                               AAA73331;
             361
                                       141
```

594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;

Sequence

```
160
                                                                                                                                                                                                                                                                                                                                                                                  480
                                                                                                                                                                                                                                                                                                                          140
                                                                                                                                                                                                                                                                                                                                            420
                                                                                                                                                                                                                                                                                                                                                                                                      161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                        540
                                                                                                                                                                 101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                            40
                                                                                                                                                                                                         80
                                                                                                                                                                                                                                                                                                                       61 ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                           181 ACCTGCATCAACGGTGTTTGCTGGACCGTTTACCACGGTGCTGGTACCGTACCATCGCT
                                                                                                                            21 TyralaginginthrargglygluginglycysglnLysthrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AlavalAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3
                                                                                                                                                                                                                                                                                                                                                                                                                                                       /*tag- a
/product- "NS4A-NS3 fusion protein #3"
594
191
2
0
0
Length:
Matches:
Conservative:
Mismatches:
                                      Indels:
                                                Gaps:
                                                                   (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                   US-09-965-594-20 (1-197) x AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAA73330 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
```

S

```
protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helixO wild-type sequence.
161 ProAlaGlyHisAlaValGlyIlePheArgAlaValSerThrArgGlyValAlaLyS 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4 fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                          protease; viral replication; chronic liver disease; liver cancer; mutant; mutein; ds.
                                                                                                                                181 AlavalaspPheIleProvalGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8
                                                                                                                                                   /*tag= a
/product= "NS4A-NS3 fusion protein #8"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6.41e-80
946.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271.
                                                                                                                                                                                                                                            AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-465976/40.
P-PSDB; AAB15226.
                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                  liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                 AAA73335;
                                                                                                                                                                                                       AAA73335
                                                                                                                                                                                                                                            The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the POlyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver fallure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61 TACGCTCAGCAGACTCGAGGTGAGGGTTGCCAAGAAACCTCCCAGACGACGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening Inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21 TyralaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 594 BP; 103 A; 186 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          594
1188
7
0
                                                                                                                                  Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-20 (1-197) x AAA73330 (1-594)
                                                                                                                                  Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                     Claim 26; Fig 13; 66pp; English.
                                                                                              (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1.87e-82
973.00
96.45%
95.43%
                     06-JAN-2000; 2000WO-US00345.
                                                         990S-0115271
                                                                                                                                                                 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                                                       P-PSDB; AAB15221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                         08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
```

g ò a ö

ð

õ g ò å ö 셤 ò ø

```
420
                                                                                                                                                                                                                                                                                                                          540
                                                                                                                                                          180
                                                                                                                                                                                              240
                                                                                                                                                                                                               100
                                                                                                                                                                                                                                                  120
                                                                                                                                                                                                                                                                     360
                                                                                                                                                                                                                                                                                       140
                                                                                                                                                                                                                                                                                                                                                                ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                 300
                                                                                                              9
                                                                                                                                                                            80
                                                                                                    40
                                                                                  9
                                                                                                                                                                                                                         301 CCGCTCCGCAGGGTTCCCGTTCCCTGACCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                   GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                       LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                        TyralaGlnGlnThrargGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                        LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                             ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                              ACCIGCATCAACGGIGTTIGCIGGACCGTTIACCACGGIGCIGGIACCGGTACCAICGCI
                                                                MetLysiysLysGlySerValVallleValGlyArgIlcAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication: chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                    AlaValAspPheIleProValGluScrLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                              /*tag= a
/product= "NS4A-NS3 fusion protein #2"
0000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease coding
        Mismatches:
                   Indels:
                                               (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                               US-09-965-594-20 (1-197) x AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
94.42%
94.42%
92.75%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
Synthetic.
          Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       W0200040707-A1
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000
                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                         361
                                                                                                                                                                                                                                                                                                                                             421
                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                    21
                                                                                                                                         41
                                                                                                                                                                             61
                                                                                                                                                                                                                                                   101
                                                                                                                                                                                                                                                                                       121
                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                               161
                                                                                                                                                                                                                                                                                                                                                                                 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAA73329;
                                                                                                                                                                                                                81
                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                    ò
                                                                                                                      g
                                                                                                                                      ò
                                                                                                                                                        qq
                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                 ð
                                                                                                                                                                                          a
                                                                                                                                                                                                            à
                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                               ò
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          141 LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  295 ccGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            54
                                                                                                                                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21 TyralaglnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SerProLysGlyProValThrGlnMetLyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        588
184
3
2
                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-20 (1-197) x AAA73329 (1-588)
                                                                         2hang Y,
                                           ဗွ
                                                                                                                                                                                                                              Claim 26; Fig 12; 66pp; English.
                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2.88e-79
939.00
94.92%
93.40%
                                                                         Weinheimer S,
             99US-0115271
                                                                                                      WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                     P-PSDB; AAB15220
                                                                         Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
           08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ..
Q
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

~

AAA73328

RESULT 셤

g

```
LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                              21 TyrAlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                   41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pestivirus; Npro; protease; NS3; screening; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Chimeric - Bovine viral diarrhea virus
 US-09-965-594-20 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chimeric BVDV/HCV NS3-wt sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ABA95615 standard; DNA; 12734
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chimeric - Repatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ran
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Lai VCH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  04 - DEC - 2001.
                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                      355
                                                                                                                                                                                                                                                                                                                                                                                                141
                                                                                                                                                                                                                                                                                                                                                                                                                         415
                                                                                                                                                                                                                                                                                                                                                                                                                                                   161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              535
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABA95615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hong 2,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 9
ABA95615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                  a
                                                                            ò
                                                                                                     đ
                                                                                                                                 ò
                                                                                                                                                        g
                                                                                                                                                                                  ò
                                                                                                                                                                                                           q
                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          а
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                  Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            well as enabling structural studies of the protease and
protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            588
182
12
12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                             /*tag= a
/product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Wittekind M, Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
1..588
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Disclosure; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                 ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9.89e-77
912.00
92.89%
92.39%
89.41%
                                                                                                                                 AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                   06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0115271
                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-465976/40.
P-PSDB; AAB15212.
                                                                                                                                                                                                                                                                           Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                   19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                            13-JUL-2000
                                                  181
```

AAA73328

XX
AC
XX
AAA7
AAAA7
AAA7
AAA7
AAAA7
AAAA7
AAA7
AAAA7
AAAA7
AAAA7
AAAA7
AAAA7
AAAA7
AAAA7
AAAA7

160

234

90

0.7

9

œ

us-09-965-594-20.rng

```
The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a genome encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimeric clone of BVDV (bovine viral diarrheavirus)/HCV NS3-wt, which was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           113 GGTAGTGTTGTTGTTGTTGGTAGAATTGTTTTATCTGGTAGTGGTAGTATCACGGCGTAC 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit
                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 GlyServalvalilevalGlyArgileAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      653 CCCAAGGGTCCTGTCATCCAGATGTATACCAATGTGGACCAAGACCTTGTGGGCTGGCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                      Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMctArgScr 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-20 (1-197) x ABA95615 (1-12734)
                                                               Example 2; Columns 17-28; 20pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                               3.12e-73
892.50
92.82%
90.77%
87.50%
                                                                                                                                                                                                                                                                                                                                                                     Similarity:
                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                 HCV infection
                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         142
                                                                                                                                                                                                                                                                                                                                                                   Best Local S.
Query Match:
                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 10
요
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                윤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

AAX80355 standard; cDNA; 1998

AAX80355 ID AAX8 XX

```
243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    183
                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A perinde, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASS activity of NS3. The covalent COVAINTS COMPLEXES are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease, hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1998
168
15
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                      HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                              Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                     New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (1-1998)
                                                                                                                                                                                                                                                                                                                                                ĸ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-20 (1-197) x AAX80355
                                                                                                                                                                                                                                                                                                                                                Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1.41e-73
885.50
93.37%
85.71%
86.81%
                                                                                                                                                                                                                                          98WO-US24528
                                                                                                                                                                                                                                                                         98US-0094331
                                                                                                                                                                                                                                                                                      97US-0067315
                           (first entry)
                                                                                                                                                                                                                                                                                                                                                Taremi SS,
                                                                                                                                                                                                                                                                                                                   (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                          Hepatítis C virus.
                                                                                                                                                                                   W09928482-A2
                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                            24-NOV-1998;
                                                                                                                                                                                                                                                                         28-JUL-1998;
                         07-SEP-1999
                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                Malcolm BA,
                                                                                                                                                        Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

us-09-965-594-20.rng

```
NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASE activity of NS3. The covalent covalent covalent complexes are more soluble, stable and active than the non-
                           363
                                                      121
                                                                                423
                                                                                                           141
                                                                                                                                    483
                                                                                                                                                              LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                         543
                                                                                                                                                                                                                   603
              ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                            ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                       AlabroglnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                         ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Weber PC,
                                                                                                                                                                                                                                                                                                                                                         AAX80359 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                              07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                           304
                                                       102
                                                                                                            122
                                                                                                                                      424
                                                                                                                                                                 142
                                                                                                                                                                                         484
                                                                                                                                                                                                                                                                                                   604
                                                                                                                                                                                                                                                                                                                                                                                   AAX80359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                           οŽ
                                                                                                                                Ω
                                                                                                                                                             ò
                                                                                                                                                                                    q
                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                         oy
og
                      용
                                                    ò
```

```
141
                                                                                                                                                                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                   423
                                                                                                                                                                                                                                                                                                                                                                                         162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                             243
                                                                                                                                                                                                                                303
                                                                                                                                                                                                                                                101
                                                                                                                                                                                                                                                                  363
                                                                                                                                                                                                                                                                                                                                                       LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                 124 TCCCAACAGACGCGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG 183
                                                                                                                                         41
                                                                                                                                                                            61
                                                                                                                                                                                                               81
                                                                                                       21
                                                                                                                                                                                                                                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                               SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                      ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                  5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                            other:
                         1998
167
16
10
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104.
                                           Conservative:
Mismatches:
Indels:
0
                          Length:
Matches:
569 G; 423
                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BP.
ö
                                                                                     US-09-965-594-20 (1-197) x AAX80359
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80354 standard; cDNA; 1998
BP; 411 A; 595
                        2.69e-73
882.50
93.37%
85.20%
86.52%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus Synthetic.
                                           Percent Similarity:
Best Local Similarity:
Sequence 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    07-SEP-1999
                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                82
                                                                                                                                                                                                                                                                                                                                                        142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAX80354;
                                                                                                                                                                             42
                                                                                                                                                                                                                                                                                                                                                                        484
                                                             Query Match:
                           ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX80354
ID AAX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                       ŏ
                                                                                                                       đ
                                                                                                                                          ò
                                                                                                                                                          a
                                                                                                                                                                            ò
                                                                                                                                                                                            qq
                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                             ö
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent Covalent protease scivity stands complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
22 AlaginginthrArgglyGlugingiyCysGlnLysthrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV: hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                          Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                        HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 147-148; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-20 (1-197) x AAX80345 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weber PC,
                                                                                                                                                                                                                                                     ВP
                                                                                                                                                                                                                                                     AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.58e-73
878.50
93.33%
85.64%
86.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           980S-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                             SO CCC CCC CCC X S X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X B X
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         අ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      364 GCGCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                                                                                                                                                                             NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A Peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGGCCTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                        invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 163-166; 211pp; English.
                                                                                                                                                                                                                                                                                                                covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-20 (1-197) x AAX80354 (1-1998)
                                                                                                                                                                                                δ,
                                                                                                                                                                                                   Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.34e-73
881.50
92.86%
85.71%
  98WO-US24528
                                                        98US-0094331.
97US-0067315.
                                                                                                                                                                                                   Taremi SS,
                                                                                                                                                                                                                                                                                                             New hepatitis C virus
                                                                                                                                           (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                        28-JUL-1998;
28-NOV-1997;
24-NOV-1998;
                                                                                                                                                                                                   Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
DB:
```

651 167 15 10 3

a

a ò g ò

g õ

a

å a à

ò

g ö g ò

ö

```
AAX80353 standard;
                                                                                                                                              Best Local Similarity:
                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   07-SEP-1999
                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             604
                                                                                                                                                                                                                                                                    42
                                                                                                                                                                                                                                                                                                                                                                                                         122
                                                                                                                                                                                                                                                                                                     62
                                                                                                                                                                                                                                                                                                                                                                        102
                                                                                                                                                                                                 'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAX80353
                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAX80353
8×88888888888×8
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                           QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                 603
                                         243
                                                                           303
                                                                                                            363
                                                                                                                             121
                                                                                                                                              423
                                                                                                                                                              141
                                                                                                                                                                               483
                                                                                                                                                                                                 161
                                                                                                                                                                                                                 543
                                                                                                                                                                                                                                  181
101
                        61
                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV
                                                                                                    84 AACCAGGICGAGGAGGAGGITCAGGIGGIIICCACCGCAACACAAICCIICCIGGCGACC
                                                          62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                    244 TGCGTCAACGGCGTGTTGGACCGTTTACCAIGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                             364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG
                                                                                                                                                                                                142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                        AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                           82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                              ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A.NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                            182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                       HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Page 176-179; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ,
                                                                                                                                                                                                                                                                                                                      BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   describes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Weber
                                                                                                                                                                                                                                                                                                                    AAX80358 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             97US-0067315
                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Taremi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                      07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                             102
                                                                                                                                                              122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Malcolm
                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                  a
                                                                         q
                                                                                            à
                                                                                                        ద
                                                                                                                             ò
                                                                                                                                                                           g
                                                                                                                                                                                                                9
                                                                                                                                                                                                                                 ò
                        ò
                                       셤
                                                       ò
                                                                                                                                                              à
                                                                                                                                                                                                ç
```

```
121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81
                hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GlyServalValileValGlyArgileAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      364 GCGCCCCCCCGGGGGGGGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
peptide, a linker, and an HCV NS3 serine protease domain, where
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                       Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                  1998
167
15
11
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NO:103.
                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complex encoding cDNA SEQ ID
                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-20 (1-197) x AAX80358 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CDNA; 1998
                                                                                                                                                                                                                                                                              6.39e-73
878.50
92.86%
85.20%
86.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
```

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex. Comprising a central hydrophobic domain of native HCV NS3 earlie protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPRSE activity of NS3. The covalent covalent protease-peptide complexes previously available.
HCV; hepatitis C virus; single chain recombinant complex; linker;
                                           NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Page 160-162; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weber
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-385385/32
                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                          WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24 - NOV - 1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
                                                                                                                                                                                                                                        Synthetic
NAME OF STREET O
```

US-09-965-594-20 (1-197) x AAX80353 (1-1998)

1998 167 11 11

Length:
Matches:
Conservative:
Mismatches:
Indels:

7.93e-73 877.50 92.86% 85.20% 86.03%

> Percent Similarity: Best Local Similarity: Query Match: DB:

Alignment Scores:

Score:

22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSerHisThrGlyArgAspLys

42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61

41

243

81

AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121

102

*

244 TGCGTCAACGGCGTGTGTGTGTGCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303

```
122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                           543
                                                                                            161
364 GCGCCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423
                             182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                             Search completed: August 30, 2003, 19:48:15 Job time : 188.939 secs
                                   g
                                                       ò
                                                                       셤
                                                                                            ò
                                                                                                               q
               õ
                                                                                                                                   õ
                                                                                                                                                    a
```

```
GenCore version 5.1.6
```

Ltd.	
Compugen Ltd.	
2003	
1993	
Ω̈	
Copyright (c) 1993 - 2	

OM protein - nucleic search, using frame_plus_p2n model

August 30, 2003, 19:20:43 ; Search time 1910.31 Seconds (Without alignments) 2506.388 Willion cell updates/sec	US-09-965-594-20 score: 1020 e: 1 MKKKGSVVIVGRINLSGDTAVAKAVDFIPVESLETTMRSP 197	table: BLOSUM62 Xgapop 10.0 , Xgapoxt 0.5 Ygapop 10.0 , Ygapoxt 0.5 Fgapop 6.0 , Fgapoxt 7.0 Delop 6.0 , Delext 7.0	d: 22781392 seqs, 12152238056 residues	Total number of hits satisfying chosen parameters: 45562784
Kun on:	Title: Perfect score: Sequence:	Scoring table:	Searched:	Total number of

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Database :

gb_est4:*
gb_est5:*
em_estfun:* em_gss_hum:* em_gss_inv:* em_gss_pln:* em_gss_vrt:* em_gss_fun:* em_esthum: em_esthum: em_estin: em_estin: em_estiv: em_estip: em_trc:* qb_esti:* qb_esti:* qb_esti:* em_estom:* gb_est3:*

29: qb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

EST 29-APR-2003 EST.

Mus musculus (house mouse)

Mus musculus

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1031) CB950999 1031 bp mRNA linear EST AGENCOURT_13445496 NIH_MGC_177 Mus musculus cDNA clone IMAGE:30316162 5', mRNA sequence. CB950999.1 GI:30205777 CB950999 LOCUS ACCESSION VERSION KEYWORDS SOURCE ORGANISM RESULT 1 CB950999

REFERENCE

em_gss_pro:*
em_gss_rod:*
em_gss_phg:*
em_gss_vrl:*
gb_gssl:*

em_gss_mam:* em_gss_mus: us-09-965-594-20.rst

```
US-09-965-594-20 (1-197) x BM915803 (1-1146)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                5', mRNA sequence.
BM915803
BM915803.1 GI:19366182
                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         106.00
36.26%
25.82%
10.39%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Source
                                                                                                                                                                                                                                                                                                                                                                                                              KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ..
Q
                                                                                                                                                                                                                        RESULT 2
BM915803/c
                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                          VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                 q
                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /ullerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarellerarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelrarelra
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  124 ArgHisAlaAsp-ValIleProValArgArgArgGlyAspSerArgGlySerLeuLeuSe 143
                                                                     Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Michael Brownstein
CDNA Library Preparation: Michael Brownstein Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: MDCMIO7 row: b column: 11
High quality sequence stop: 333.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               143 rProArgProlleSerTyrLeuLysGlySerSerGly-----GlyProLeuLeuCy 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        356 CAAGTGTGCCACCTGCTCACCCAC.----AAGGTTCAGTATGIGATTCAAGGCTATCCC 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   509 ------ACACTTAACCT-CCTTGCCTGGCCTGGCATGTTGGGCAAAAAGAACCGTTT 558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---GlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThr 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            49 GlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrp 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29 GlnGlyCysGlnLysThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluVal 48
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89 MetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGln-------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCCATGGGGCTTGTTCTCCCGGCCTCTCCCGTGGGCAATTACAAAACNCCCTTAACGG
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1031
54
19
69
37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels:
                                                                                                                                                                                                                                                                                                                                                                                                              1. .1031
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   us-09-965-594-20 (1-197) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2.23
106.50
41.018
30.348
10.448
                                                   Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
ORIGIN
                       TITLE
JOURNAL
COMMENT
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
1. .1146
/organism="Homo saplens"
/mol_type="mina"
/db xref='taxon:9606"
/clone="IMAGE:5482056"
/clone="IMAGE:5482056"
/tlosue_type="amelanotic melanoma, cell line"
/tab.host="amelanotic melanoma, cell line"
/tab.host="amelanotic melanoma, cell line"
/lone="lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site_1: xho!; Site_2:
FOORI; cDNA made by oligo-dT priming. Directionally cloned
into Ecorl/xho! sites using the following 5' adaptor:
GGACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library.
8 141 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BM915803 1146 bp mRNA linear EST 12-MAR-2002 AGENCOURT_6639455 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5482056
160 sProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLy 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 AlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSer-HisThrGlyArgAspLy 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41 sAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaTh 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota: Metazoa; Chordata; Craniata: Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1146)
NH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.oih.gov
Tissue Procurement: DCTD/DTP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clound through the I.M.A.G.E. Consortium/Linx at:
http://image.lln.gov
Plate: LLCA2007 row: i column: 01
High quality sequence start: 6
High quality sequence start: 6
High quality sequence store.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 þe
                                          sAlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1146
47
19
59
57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
```

```
Contact: Robert Strausberg, Ph.D.

Email: ggapbs r@mail.nh.gov

Email: cgapbs remail.nh.gov

Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Washington University Genome Sequencing

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LiNL at:

http://image.llnl.gov
                                                                                                                                                                                                                                                                                                                                                                                      BG089727 629 bp mRNA linear EST 26-JAN-2001 mab90e06.x1 NCI_CGAP_SP2 Mus musculus cDNA clone IMAGE:3977578 3' similar to SW:GRAD_MOUSE P11033 GRANZYME D PRECURSOR;, mRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /clone='INAGE:397758"
/tissue_type="NK cells (flow-sorted)"
/tab_nost='DH10B (Tl-resistant)"
/clone_lib="NCI_CGA=Sp2"
/note="Corpus: splean: Vector: pCMV-SPORT6 (Life
/note="Corpus: splean: Vector: pCMV-SPORT6 (Life
Technologies); mRNA made from flow-sorted NK cells, cDNA
made by oligo-dT priming. Directionally cloned. Average
insert size 1.5 kb. Primary library, non-amplified. cDNA
Library Preparation: David B. Kriman, Ph.D."

156 c 150 g 191 t l others
                                                                                                                                                                                     OLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGl 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 629)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
NATIORAL Cancer Institute, Cancer Genome Anatomy Project (CGAP), Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThr 55
                                                    --- GlySerSerGlyGlyPr
                                                                                                  CCACCCTCGGCCCACCGATGCGCTAAGCCTTTTACAAGCCACGGCCCGGGGCCCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     629
48
23
59
51
                                                                                                                                                                                                                                                                                         877 CGTTTTAAAGCCCCCGGCCTTCGCGCCGGGCGAAGCA 913
                                                                                                                                                                                                                                            177 yvalAlaLysAlaValAspPheIleProValGluSer 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                    -SerProArgProlleSerTyrLeuLys----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Seq primer: -400V Irom with thigh quality sequence stop: 422.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-20 (1-197) x BG089727 (1-629)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mus musculus (house mouse)
Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BG089727.1 GI:12572290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sednence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                 157
                                                                                                                                                                                                                                                                                                                                      RESULT 5
BG089727/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                  ö
                                                                                             g
                                                                                                                                            òγ
                                                                                                                                                                                       αd
                                                                                                                                                                                                                                            Qγ
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                            /Act. marge "adult, 16 yr"
//dev_stage="adult, 16 yr"
//dev_stage="adult, 16 yr"
//dev_stage="adult, 16 yr"
//deb.nost="PHB10B"
//clone_lib="Lupski_sympathetic_trunk"
//note="Vector: pcMv-SPORT6 (Life Technologies); Site_l:
//note="Vector: pcMv-SPORT6 (Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine); available through Life
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               490 CCATTTTTAATAAAGGGTGCTCTGTTAATCATGGCGCCCACGGGCCCGCTGATACTTCCA 549
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              90 TyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeu 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      430 CCAGCCACTGTCTTCATGTAATGACCCTTTCCCACCACACAGGGCACACGCATGGGAT 489
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         370 CATGGGGAAATCGCCCCCTTCAGAGACGAGGTCGTTCATGTTTTCTGAAAACATAACCG 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyCysGlnLysThrSer 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AlaThrGlnThrPheLeu------AlaThrSerIleAsnGlyValLeuTrpThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             577 GCTGCCTGCACACCCCCCATGACCGATGTTGGGCTTATGTGGAAACGGCGAGGCGTTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -----ArgArgGlyAspSerArgGlySerLeuLeu-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      70 ValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITACCACATGIGACAGTGACITI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr-----
                                                                             ည္ရ
                 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can http://image.lln.gov
http://image.lln.gov
Plate: LiAM13595 row: column: 13
High quality sequence start: 57
High quality sequence stop: 394.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             121 -LeuValThr------ArgHisAlaAspValIleProValArg----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          __::::===:::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28 others
cDNA Library Preparation: Life Technologies, Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      211 t
                                                                                                                                                                                                                                                              /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6192708"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-20 (1-197) x BQ892487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      343 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rechnologies.
362 c
                                                                                                                                                                                                                                                                                                                                                                /sex-"male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9.59
101.00
34.74%
23.47%
9.90%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ..
2
                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
```

õ đ ò qq δž g ò a ò අ à 셤 ç g

```
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ..
Э
   ORGANISM
                                                    REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 9
AQ538021
                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                              /#INTALCIAND wild type mt+ 21gr"
/#INTALCIAND wild type mt+ 21gr"
// db_xref="taxon:3055"
// clone=lib="Cc. reinhardii CC-1690, Stress condition I,
normalized, Lambda Zap II"
// note="vector: pBluescript II SK-; Site_I: ECORI; Site_2:
XhoI; This library, constructed by John Davies and Jeffrey
MCDermott, combines cDNAs from CC-1690 cells grown to
mid-log phase in TAP-N (30 min, 1hr, 4hr), TAP-S (30 min,
1hr, 4hr), TAP-P (4hr, 12hr, 24hr), No3 to NH4 (30min, 1hr,
4hr) and NH4 to NO3 (30min, 1hr, 4hr). PolyA mRNA was
purified from each sample, pooled and cDNA synthesized.
The CDNA was directionally cloned into lambda Zap II
(Stratagene) in the ECORI (5') and XhoRI (3') sites.
pBluescript II SK- plasmids were excised from the lambda
ZAP clones by superinfection with ExAssist (Stratagene)
phage: The library was normalized using method 4 described
in Bonaldo et al (1996) Genome Research 6: 791-806."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                brjoubby 984 bp mRNA linear EST 21-NOV-2000 6018888252F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4122276 5', mana saguence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       405 GAAGGGGACGGGAATICCGTTTGCGGAAAAGTGAGCGCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    150 LeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePhe 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           107 ArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       170 ArgAlaAlaVal--------------SerThrArgGlyValAlaLySAla--- 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     465 CTCCAAGGCAGCAATGGGAGCCTTTCGCGGTGTGCGGGTGCTCGTCCTCTAATGTCAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----GlySerLeuLeuSerProArgProIleSer---Tyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAGATCATTTGCATGTGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTTATAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      525 AAAGAAACCATTTGAGTAGGAAGTCGCCCGTTTACCCCCGGAAGGTGAAGGTCACCTCTAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          231 ACAAAGTTACATACACGG-----AAGGACCAGCGCGCTTGGCCCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182 ValAspPheIleProValGluSerLeuGluThrThrMetArg 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ||{||{
Grggaccgcatacatctggaagacacacggtgcactctacga 626
'organism="Chlamydomonas reinhardtii"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71 TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys----
                                                                                                                                                                                                                                                                                                                                                                                                  701
38
22
22
63
31
                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-20 (1-197) x BFB63244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BF304699.1 GI:11251586
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                16.8
95.50
38.96%
24.68%
9.36%
                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BF304699
                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 8
BF304699/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                    ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
```

```
/close="INAGES:4122276"
/(15sue_type="Thabdomyosarcoma"
/lab_host=="bhi0b" (phage-resistant)"
/clone_lib="NHLMGC_l7"
/clone_lib="NHLMGC_l
                                                              in (bases 1 to 98)

NIH-MGC http://mgc.nci.nlh.gov/.

NIH-MGC http://mgc.nci.nlh.gov/.

NIH-MGC http://mgc.nci.nlh.gov/.

Upublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ArCc
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ArCc
Contact: Nobert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ArCc
Contact: Nobert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ArCc
Contact: Arcayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: The I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCM1005 row: g column: 13
High quality sequence stop: 646.

Location/Qualifiers

rcc

1...984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       159 LeuCysProAlaGlyHis-Ala------ValGlyIlePheArg-----AlaAlaVa 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  100 TrpGlnAlaProGlnGlySerArgSerLeuThr -- ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  554
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 GlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeu 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   526 TGGTGTCCAAGCTCGCGCTGCTGTACTGGAAGTCGGCACGCTCCGGTCAGGTGCAGCTTC 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         173 ISerThrArgGlyValAlaLySAlaValAspPheIleProValGluSerLeuGluThrTh 193
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        119 LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           646 regeceagrecacecarrecergesaaaaagacerergee----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    984
66
26
34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-20 (1-197) x BF304699 (1-984)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  466 CAGCGCCCGGGGTGCG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                193 rMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CAAGGGGGCGCCC 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26.2
95.50
42.86%
37.14%
9.36%
```

```
BE868757 10-0CT-2000 01445993F1 NIH_MGC_65 Homo sapiens cDNA clone IMAGE:3850079 5',
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="Homo sapiens"
/mol_type="mxNA"
/db_xref="taxon:9606"
/clone="IMAGE:3950079"
/tissue_type="adenocarcinoma"
/tissue_t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ...--ArgSerLeuThrProCysThrCysGlySerSerAspLeuTy 120
                                                                                                                                          120 rLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySe 140
                                                                                                                                                                                                                                                      140 rLeuLeuSerProArgProIleSerTyrLeuLysGlySer-SerGlyGlyProLeuLeuC 160
                                                                                                                                                                                                                                                                                                             361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MetLysLysLysGlySerValValIleValGly-----ArgIleAsnLeuSerGly 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Butheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 935) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM9568 row: i column: 24
High quality sequence stop: 557.
Location/Oualifiers
1. 935
                                                                                                                                                                                267 -------CTGGAGCAGATATGACGAGCCATGGAGGAGGAGAGCCT
                                                                                                                                                                                                                                                                                                             311 TCTCCCTGGCCCAGCCCTGGACCATATCTGAAAGCTGTGAAGTGGAGGC-----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               935
117
128
58
6
                                                                                                                                                                                                                                                                                                                                                                    160 ysProAlaGlyHisAlaValGlyIlePheArgAlaAla 172
                                                                                                                                                                                                                                                                                                                                                                                                                        362 GCCGGCGGGGAGCTGATGCAGCTGTTCTCACAGGGGCT 399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
                                                                  145 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: ATCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-20 (1-197) x BE868757 (1-935)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 261 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BE868757.1 GI:10317533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Technologies.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34.5
94.00
42.668
30.778
9.228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 245 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             mRNA sequence.
BE868757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ĸ
                                  106 r----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
BE868757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                      셤
                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                             ò
560 bp DNA linear GSS 18-MAY-1999
RPCT-11-32014.TJ RPCT-11 Homo sapiens genomic clone RPCT-11-32014,
appendic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            87 ThrGlnMetTyrThrAsn-ValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySe 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
Clones are derived from the human BAC library availability, please contact Pieter de Jong
(pieter@dejong med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
thtp://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  128 AIGIGG------CACGGATITGGAGCCAGGATICCAGCIGGCICIAAAGGG----- 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          86 TICTGCCAGTCTCTGTCATCCACTGGCCAAGAG---------AATTGCAAA 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GlnGlyCysGlnLys-----ThrSerHisThrGlyArgAspLysAsnGlnValGluGly 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26 CAAIGCTGTGAGAACCCAGTGGATGGACACCCTGGAAGGGCTATGAGATCATAGGAGGGC 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   47 GluvalGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal 66
                                                                                                                                                                                                                                          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 560)
                                                                                                                                                                                                                                                                                                                                       Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
                                                                                                                                                                                                                                                                                                                                                                                              Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /cell_rype="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCII Human Male BAC Library"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67 LeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                            Unpublished
Other GSSs: RPC1-11-32014.TV
Other GSSs: RPC1-11-32014.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Fax: 301 838 0208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     560
112
138
38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism-"Homo sapiens"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /mol_type="genomic_DNA"
/db_xref="GDB:7622691"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /db_xref="taxon:9606"
/clone="RPCI-11-32014"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-20 (1-197) x AQ538021 (1-560)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualiflers
1. .560
                                                                                                                                          AQ538021.1 GI:4849711
                                                                                                                                                                                                 Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  17.7
94.00
37.668
29.878
9.228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ends
                                                                                                                                                                                                                                                                                                                                                                                                                           Map Building
                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ø
                                                                                                                   AQ53802]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                     GSS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                          VERSION
KEYWORDS
SOURCE
ORGANISM
                               LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
COMMENT
                                                                                                                ACCESSION
                                                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                              TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
```

FEATURES the other source misc_feature	117 BASE COUNT 274 a 264 c 242 g	00.001 QY 32 GlnLysThrSerHisThrGly QY 11	Oy 74 AlaciyThrargThrileAlaSerPro 1	Oy 114 CysGlySerSerAspLeuTyrLeuVal	Oy 154 Ser
18 ASPINIAGATATAGATAGATAGAGATAGAGATAGAGATAGAGATAGATAGATAGATAGATAGAGAGAGAGAGAGAGAGAGAGAGAGATAGATATATAGATATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATATAGATA	98 ValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSer 1	CNSO6OHN T3 end of clone AWOAAO06B03 of library AWOAA from strain CLIB 89 of Yarrowla lipolytica, genomic survey sequence. AL410673. AL410673.1 GI:12179275 GSS. Yarrowla lipolytica Yarrowla lipolytica Yarrowla lipolytica Surrowla lipolytica Surrowla lipolytica Yarrowla lipolytica	Suciet, J. Aigle M., Artiguenave, F., Blandin, G., Souciet, J., Aigle M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brottier, P., Gasaregola, S., Belottin-Fukuhara, M., Duyon, B., Durrans, P., Lepinqte, A., Llorente, B., Malpertuy, A., Neuvelise, C., Ozier-Kalogeropoulos, O., Potier, S., Saurin, W., Tekaia, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Wincker, P. and Weissenbach, J. Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies FEBS Lett. 487 (1), 3-12 (2000) 11152876	a, S., tve, F., txplora	3 (bases 1 to 1062) 3 (bases 1 to 1062) Genoscope. Direct Submission Submitted (07-SEP-2000) Genoscope - Centrc National de Sequencage, 2 rue Gaston Crembeux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: 9 seqref@enoscope.cns.fr - Web: www.genoscope.cns.fr) This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii. Zygosaccharomyces rouxii. Saccharomyces kluyver: Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
3 4 3 4 3 4 3 4	oy Oy Oy	RESULT 11 CNSOGHN LOCUS DEFINITION ACCESSION REXTWORDS SOURCE ORGANISM	AUTHORS TITLE JOURNAL MEDLINE PUBMED	REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED	REFERENCE AUTHORS TITLE JOURNAL COMMENT

```
GAATCATAGAGCACCAGAAAGCTGCTCTAGGAATC 1013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ::: ||||
SATACTCTCAGACA-----TGG 548
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CCACGGGCTCAGAATTGCCACAGACGAAGATCGA 608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             111
3CTGGGGAGCCACACCTGCCGGTTCTCAGTTGACA 788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    alThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --GlyGlyProLeuLeuCysProAlaGlyHis--- 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ----SerLeuThrProCys-----Thr 113
                                                                                                                                                                                                                                                                                                                                                                                       nrPheLeuAlaThrSerIleAsnGlyValLeuTrp 68
                                                                                                                                                                                                                                                                                                                                                                                                                                    73
                                                                                                                                                      Saccharomyces cerevisiae ORF YKR093w [
sporter ]"
:imental
                                                                                                                                                                                                                                                                                                                                          yArgAspLysAsnGlnValGlu---GlyGluVal 48
extremities were sequenced. See this sequence and for the sequence of insert.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                oLysGlyProValThrGlnMet------
                                                                                                                                                                                         2 others
                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                               lipolytica"
DNA"
                                                                                                                                                                                           280 t
                                                                                                 .
.
```

6

```
Eukaryota; Viridiplantae; Streptophyta; Embryophyta: Tracheophyta; Spermatophyta; Tracheophyta; Spermatophyta; Tracheophyta; Spermatophyta; Paulcoldeae; Paulcoldeae; Andropogoneae; Zea.

1 (bases 1 to 789)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick Maize Genomics Consortium
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------11ePro 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Zea mays"
/mol_type="genomic DNA"
/strain="B13"
/db_xref="taxon:4577"
/clone="Inh="ZM_0:6_1.0_KB"
/note="Vector: pCK4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"
a 220 c 203 g 156 t
                      ArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyPro 157
                                                                                   158 LeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGly 177
                                                                                                                                                                                   244 ------GCCATTGACATCAACAAGGAGAAGTTCACCATTGAGGCATATGCAATTCCC 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     48 ValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal--- 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29 GlnGlyCysGlnLysThrSerHisThrGlyArgAspLysAsnGlnVal---GluGlyGlu 47
                                                                                                        ::: |||
274 GITITCTGCCAG--------GGAGIAACTCGTCGCGCA------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ::::::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        789
449
467
100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches:
Conservative:
Mismatches:
                                                                                                                                                 178 ValAlaLysAlaValAspPhe-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ::
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-20 (1-197) x CC406705 (1-789)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Class: sheared ends.
Location/Qualifiers
1. .789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Email: whitelaw@tigr.org
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Cathy Whitelaw
                                                                                                                                                                                                                                        193 TIGGATACATTIGAG 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Other_GSSs: PUHKD12TB
                                                                                                                                                                                                                  187 ValGluSerLeuGlu 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93.50
39.46%
26.49%
9.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seq primer: TF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                 Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TIGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No.:
                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                                                                                                              CC406705
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                              đ
                                                                                                              g
                                                                                                                                               ô
                                                                                                                                                                                 qq
                                                                                                                                                                                                                  à
                                                                                   οy
          772 bp DNA linear GSS 19-MAY-2003 genomic survey sequence.
                                                                                                                                                  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoldeae; Andropogoneae; Zea.

1 (bases 1 to 772)
Whitelaw.C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennetzen,J.
Maize Genomics Consortium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :::|||
549 ATCTGGAGG------TCACAGACTTTGCTGACGATGAGATCAAGGGTG 508
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     87 ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp------GlnAlaPro 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACACAGATGATGACCAGG-----AAGAGCAGGACCCAGTGGTTTCCCTACATGCAATAA 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu------ 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   453 CTGGGATTAGAAGGGAGGACACCATGCAGCTGCGGGTAGTCCTCAACGGTCAGGAGCTGC 394
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -----TyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSer 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   393 TGGCCCTACTIGACACGGGTTCAACACATAACTTCATCAACTGCAAGGCGGCA-CAACAG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /db_xref-*taxon:4577*
/clone="ZMBTa469B24"
/clone_lib="ZM_0.6_1.0_xB"
/clone_lib="XM_0.6_1.0_xB"
/once="Vector: pCR4 -TOPO; Site_1; EcoRI; 0.6-1.0 kb high
COT selected genomic DNA library"
198 c 216 g 208 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      29 GlnGlyCysGlnLysThrSerHisThrGlyArgAspLysAsnGlnVal ---GluGlyGlu 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    48 ValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  603 GTTACAATTGTGACG-----AGCCCTATGTACGAGGCCACCGCTGTCCCAGGCTATTCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal
                                                                                                                                                                                                                                                                                                                                           9712 Medical Center Drive, Rockville, MD 20850, USA 121: 301-888-5643 Fax: 301-888-5643 Email: whitelaw@tigr.org Seq primer: TR Class: sheared ends. Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        772
49
24
67
67
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:

    772 / Organism-"Zea mays" / Mol_type-"genomic DNA" / Strain-"B73"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-20 (1-197) x CC406704 (1-772)
                                                                                                 CC406704.1 GI:30886794
GSS.
                                                                                                                                                                                                                                                                                                               Contact: Cathy Whitelaw
                                                                                                                                                                                                                                                                                              Other_GSSs: PUHKD12TD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30.1
93.50
39.46%
26.49%
9.17%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                 Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                 TIGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                 DEFINITION
                                                                                                                                                 ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                  CC406704/c
                                                                                                                                                                                                                                                            TITLE
JOURNAL
COMMENT
                                                                                 ACCESSION
RESULT 12
                                                                                                 VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                  REFERENCE
                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

2.5 h" /dev_stage="adult" /dev_stage="adult" /clone_lib="An expressed sequence tag (EST) collection from the resurrection plant Selaginella lepidophylla" /note="vector: Lembda Uni-Zap XR, Bluescript SK-; Site_l:	Oy 94 AsplysAspleuValGlyTrpGlnAlaProClnGlySerArgSerLeuThrProCysThr 113	Db 173 GGATTGGATCATACCTGACCAGGGGGTCATCGAGGGGGATTACT 229 29 GlySerLeuLeuSerProArgProlleSerTyrLeu	SULT 15 SULT 15 SULT 15 SUS 242381/C CUS 108351 CESTION 108351 CESTION BZ3423 WHORDS GSS. WORDS GSS. WORDS GSS. CHACK SOUTHLE GENERALY SULAWE TITLE GENOMI GOORD TITLE GENOMI GOORD TITLE GENOMI CONTACT COLLAR COLL
Db 170 GTTACAATTGTGACGAGCCCTATGTACGAGGCCACCCCCAGGCTATTCT 223 Qy	LeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGly :::	2	ACTHORS Ituridad.G. and Cushman,J.C. TILE An expressed sequence tag (EST) collection from the resurrection plant Selaginella lepidophylla JOURNAL Unpublished COMMENT Contact: Cushman JC Contact: Cushman JC Copportment of Biochemistry University of Nevada MS200, Reno, NV 89557-0014, USA Tel. 775-784-1918 Fax: 775-784-1650 Email: joushman@unr.edu FRAX: 775-784-1050 Email: joushman@unr.edu FCARMARD: T7 21mer FORWARD: T7 21mer Plate: 005 row: F column: 12 Seq primer: T3 20mer BACKWARD: T7 21mer Plate: 005 row: F column: 12 Seq primer: T3 20mer High quality sequence stop: 528. Location/Qualifiers Source //db xref="taxon:59777" //clone="sla00SP12" //clone="sla00SP12" //clone="sla00SP12" //clone="sla00SP12"

```
150 LeuLysGly------LeuCy 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  160 sProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyVal---- 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----CysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     261
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           119 LeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAsp----- 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 260 GTGGAGAGGGCAGGACGACGACGA-------CGACGACGGGGTGGAGACCCG 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----SerArgGlySerLeuLeuSerProArgProIleSerTyr 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     215 GTTTGACGCGGCGCGCGCGCGCGCGCACACTCTCGCTCGGCCGTGGGAATAT 156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GCCCATGCCGTGCGGGGTGTTTTTTGAACGTACCGCACGGCGCGGGGAGGACGACAGGC 441
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GIGGCCGIGGCCCGCCGCCGCCGCGCGTGGGAATTAIGGCCGCAAIAAACTCGGCAC 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           96
                                                                                                                                                                                                                                                                                                                                                                                                                       AlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThr ---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----IleAlaSerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LeuValGlyTrpGlnAlaProGlnGlySerArg---SerLeuThrProCysThr-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    320 CACCCAGTGCCCCCGGCGGCCGCCGTACGGGGACAGGGAAACTGGTGGTGAGGATCAC
                                                                                                                                                                                                                                                                                                                701
48
14
56
51
8
                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                     Email: mccomble@cshl.org
Plate: ic83 row: b column: ll
Seq primer: -21Ml3UnivFwd
Class: shotgun
High quality sequence stop: 701.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -AlaLysAlaValAspPheIlePro 186
                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-20 (1-197) x BZ342381 (1-701)
                                                                                                                                                                                                                                                                                                                29.8
93.00
36.69%
28.40%
9.12%
Tel: 516 367 8884
Fax: 516 367 8874
                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                      108
                                                                                                                                                                                                                                                                                                       Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          440
                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                             500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        137
                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                              source
                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                 FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
```

47 rececrecerrecrecerrece 23

Search completed: August 31, 2003, 04:27:46 Job time : 1915.31 secs

us-09-965-594-22.rag

```
AAB15225 standard; protein; 197 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; AAA73334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                               August 30, 2003, 17:42:58 ; Search time 44.6227 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1: /SIDSI/goddata/geneseq/geneseqp-embl/AA1980.DAI:*
2: /SIDSI/goddata/geneseq/geneseqp-embl/AA1981.DAI:*
3: /SIDSI/goddata/geneseqy-embl/AA1981.DAI:*
4: /SIDSI/goddata/geneseqy-embl/AA1981.DAI:*
5: /SIDSI/goddata/geneseqy-embl/AA1981.DAI:*
5: /SIDSI/goddata/geneseqy-embl/AA1981.DAI:*
6: /SIDSI/goddata/geneseqy-embl/AA1985.DAI:*
7: /SIDSI/goddata/geneseqy-embl/AA1985.DAI:*
7: /SIDSI/goddata/geneseqy-embl/AA1986.DAI:*
8: /SIDSI/goddata/geneseqy-embl/AA1980.DAI:*
9: /SIDSI/goddata/geneseqy-embl/AA1981.DAI:*
11: /SIDSI/goddata/geneseqy-embl/AA1991.DAI:*
12: /SIDSI/goddata/geneseqy-embl/AA1991.DAI:*
13: /SIDSI/goddata/geneseqy-embl/AA1991.DAI:*
14: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
15: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
16: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
17: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
18: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
19: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
10: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
11: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
12: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
13: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
14: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
15: /SIDSI/goddata/geneseqy-embl/AA1993.DAI:*
                                                                                                                                                                                                                                                                               1016
1 MKKRGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /SIDS1/gcgdata/geneseq/geneseqp.embl/AA1997.DAT:*/SIDS1/gcgdata/geneseq/geneseqp.embl/AA1998.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /SIDS1/gcgdata/geneseg/genesegp-embl/AA2000.DAT:*
/SIDS1/gcgdata/geneseg/genesegp-embl/AA2001.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /SIDS1/gogdata/geneseg/genesegp-emb1/AA2002.DAT:*/SIDS1/gogdata/geneseg/genesegp-emb1/AA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                    1107863 seqs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15224
AAB15223
AAB15222
AAB15221
AAB15221
AAB15220
AAB15220
AAB15220
AAY14943
                                                                                                       OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAB15225
                                                                                                                                                                                                                                                                                                                                                          BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ü
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                       US-09-965-594-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
                                                                                                                                                                                                                                                                                  Perfect score:
                                                                                                                                                                                                                                                                                                                                                            Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                       Searched:
                                                                                                                                                                                                                                                                                                               Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database
                                                                                                                                                       Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .
9
```

NS4A-NS3	NS4A-	NS4A-NS3	HCV NS4A-NS3 COMPI	~	•	NS4A-NS3	m	HCV NS4A-NS3 COMP1	~	HCV NS4A-NS3 COMP1	HCV NS4A-NS3 COMPI	HCV NS4A-NS3 COMP1	m	~	NS4A-NS3	NS4A-NS3	NS4A-NS3	æ	HCV NS4A-NS3 COMPL		NS4A-NS3	HCV NS3 protein.	Hepatitis C virus	Hepatitis C virus	Amino acid sequenc	Hepatitis C virus	HCV genomic amino	phcvi50-encoded se	phcv176-encoded se	Seguence encoded i	sednenc	Peptide encoded by	encode	encoded	Peptide encoded by	
AAY 24947	AAY24942	AAY17880	AAY24946	AAY24941	AAY17884	AAY17879	AAY24945	AAY24940	AAY24948	AAY17883	AAY17878	AAY24944	AAY24949	AAY17890	AAY17882	AAY17886	AAY17877	AAY17887	AAY17881	AAY17885	AAY17888	AAW93482	AAY44728	AAW77397	ABP71460	AAU99289	AAR40120	AAR79223	AAR79221	AAP92041	AAP90158	AAP90164	AAP92047	AAP92050	AAP90288	
20	20	50	20	50	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	50	20	51	19	74	23	7	16	16	10	10	10	10	10	10	
665	665	216	665	665	216	216	665	665	671	216	216	999	671	215	216	216	216	215	216	216	213	631	191	3011	3011	3012	3011	687	1648	1766	1786	2261	2301	2436	2436	
85.9											٠.	84.7			٠.	•			•		83.6											•				
872.5	-															857.5	56.	85	853.5	53,	849	84	48.	æ,	48.	48.	\$ 5.	44	44	44	44	44	44	44	4.4	
10	7	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	53	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	4.5	

ALIGNMENTS

```
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1
                                                                                             Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb V;
                                              Hepatitis C virus NS4A-NS3 fusion protease #7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                               06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                08-JAN-1999; 99US-0115271
19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2000-465976/40.
                                                                                                                                                                           Hepatitis C virus
Synthetic.
```

~

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the mutaviral treatments of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helixo-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLRGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                        1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQRTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                  61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis; NS3 protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                                         0;
                                                                                                                                                                                                                                                                                                                          100.0%; Score 1016; DB 21; Length 197; 100.0%; Pred. No. 2.1e-98;
                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus NS4A-NS3 fusion protease #6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         liver failure; liver cancer; mutant; mutein
                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2
                                                               Claim 23; Fig 17; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15224 Standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                       al Similarity
197; Conserv
                                                                                                                                                                                                                                                                                           197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-PSDB; AAA73333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic
                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15224;
                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15224
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

.; 0

Gaps

120

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirthosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and proteins view of the protease and
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                        LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                             9
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                   1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                          61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                           21; Length 197;
                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease #5.
                                                                                                                                                                                                                                                                                                                      Score 1010; DB
Pred. No. 9e-98;
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               $
                                                                         Claim 23; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAB15223 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                         99.4%;
99.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                      Matches 196; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                              197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Repatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAA73332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W0200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         38-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic.
                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15223;
                                                                                                                                                                                                                                                                                                                         Ouery Match
                                                                                                                                                                                                                                                                    variant.
                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15223
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ő
```

m

```
23; Fig 15; 66pp; English.
                                                                                                                                                             AAB15222 standard; protein; 197
                                                                                                                                                                           19-DEC-2000 (first entry)
                                                                                     Matches 193; Conservative
                                                                                   Local Similarity
                                                                        197 AA:
                                                                                                                                                                                                    Hepatitis C virus
                                                                                                                                                                                                              WO200040707-A1
                                                                                                                                                                                                                            06-JAN-2000;
                                                                                                                                                                                                                     13-JUL-2000
                                                                                                                                                                                                        Synthetic.
                                                                                                           61
                                                                                                                                                                    AAB15222;
                                                                        Sequence
                                                                                                                                     181
                                                                                                                                            181
                                                                               Query Match
                     Claim
                                                                                                                                                              g
                                                                                                          ç
                                                                                                               g
                                                                                                                        ò
                                                                                                                              qq
                                                                                                                                     õ
                                                                                                                                           ద
```

08-JAN-1999;

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C viris (HCV) NG3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NG3 mutants and NG3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\cal C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis; NS3 protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    96.5%; Score 980; DB 21;
96.4%; Pred. No. 1.3e-94;
11ve 2; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease #3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  liver fallure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Wittekind M, Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15221 standard; protein; 197 AA
                                                                                                                                              Claim 23; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         181 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best_Local Similarity 96.4
Matches 190; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                             197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
N-PSDB; AAA73331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                     variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                           The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the try genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
               Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               97.9%; Score 995; DB 21; 98.0%; Pred. No. 3.4e-96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease #4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Z
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AUDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000WO-DS00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-465976/40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Wittekind M,
```

Gaps

ö

```
RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                          윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                ö
                                                                                                       The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HVV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the mutuvial treatments of the two proteins should be effective as antiviral treatments of the two proteins should be effective as antiviral treatments of the two proteins should be effective as chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                             TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                            121 LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                 1 MKKKGSVVIVGRINLSGDTAYAQQTRGEGCQETSQTGRDKNQVEGEVQIVSTAAQTFLA 60
                                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                       1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                ;
                                                                                                                                                                                                                                                                          Score 963; DB 21; Length 197;
Pred. No. 7.8e-93;
                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Weinheimer S, Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                ..
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus NS4A-NS3 fusion protease #8
                                                                                                                                                                                                                                                                                     Pred. No. 7.8e
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                       Claim 23; Fig 13; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15226 standard; protein; 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                       AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                          94.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                     Best Local Similarity 94.9
Matches 187; Conservative
2000-465976/40
                                                                                                                                                                                                                                                     197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus
           N-PSDB; AAA73330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15226;
                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                  variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                      5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                       ö
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A well as enabling structural studies of the protease and proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease: Inhibitor complexes. This sequence contains the alpha-helix0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 TSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 936; DB 21; Length 197;
Pred. No. 5.4e-90;
0; Mismatches 12; Indels (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus NS4A-NS3 fusion protease #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Æ
                                                                                                                                                                                                                                          Example 5; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB15220 standard; protein; 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 93.9%;
Matches 185; Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-DS00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000 (first entry)
WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           wild-type sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     197 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                N-PSDB; AAA73335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB15220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
```

```
N-PSDB; AAA73328
                     Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY24943;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY 24943
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TSINGVLWTVYHGAGTRT1ASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGCPLLCPAGHAVGIFRAAVSTRGVAK 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer.
                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  91.4%; Score 929; DB 21; Length 195; 92.9%; Pred. No. 2.9e-89;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
Zhang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB15212 standard: protein; 195 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                             Claim 23; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AVDFIPVESLETTMRSP 195
Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 183; Conservative
                                         WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        W0200040707-A1
  Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     variant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB1521
NAME OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
61 TSINGVLWTVYHGAGTRTIASPRGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLY 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28
                                                                                                                                                                                                                                                                                                                                                                                       The present sequence is a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replication of the virus, acting to cleave its replicative proteins from the Polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutents and NS3-MCS4 fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 MKKKGSVVIVGRIVLNG--AYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 195;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13; Indels
Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  88.8%; Score 902; DB 21; 91.9%; Pred. No. 2e-86;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1; Mismatches
      Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAY24943 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV NS4A-NS3 complex SEQ ID NO:14.
                                                                                                                                                                                                                                                                                                                                Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-DS24528
   Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 181; Conservative
                                                             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
```

9

```
122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                NSAA-NS3 complex comprising a central hydrophobic domain of native HCV hydrophobic domain of native HCV hydrophobic domain of native HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NSAA-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NSAA-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                   describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4a cofactor; NS4a-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 85.9%; Score 872.5; DB 20; Length 665; Best Local Similarity 84.7%; Pred. No. 1.4e-82; Adtches 166; Conservative 16; Mismatches 11; Indels 3;
                                                                                                                            New hepatitis C virus covalent complexes
                                                                                                                                                                   Claim 6; Page 100-102; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY24942 standard; Protein; 665 AA.
                                           Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A-NS3 complex SEQ ID NO:13.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               202 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98WO-0S24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-SEP-1999 (first entry)
                                           Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                 present invention
(SCHE ) SCHERING CORP
                                                                                    WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24 - NOV - 1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAY24942;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24942
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       62 SINGVLWTVYHGAGTRTIASPRGPVT@MYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                              The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     86.2%; Score 875.5; DB 20
85.2%; Pred. No. 6.8e-83;
tive 15; Mismatches 11
                                                                 Yao
                                                                                                                                                 New hepatitis C virus covalent complexes
                                                                                                                                                                                        Claim 6; Page 90-92; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAY24947 standard; Protein; 665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex SEQ ID NO:18.
                                                               Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 85.2%.
Matches 167; Conscrvative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                               Taremi SS,
                                                                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                               Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24947;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
```

셤

g ò

ò

qq

ò g

à

121

61 81

3;

~

```
Taremi SS,
                                                                                                                                                                                                                   216 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Malcolm BA,
 Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY24946;
                                                                                                                                                                                                                                                                                                                        62
                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 13
                                                                                                                                                                                      The
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY24946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                       â
                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                            142 VTRHADVIPVRRRGDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKA 201
                                                                                        The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-PBF10de, a linker, and an HCV NS3 Serine procease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                          VTRHADVIPVRRRGDSRCSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                          61
                                                                                                                                                                                                                                                                                                              8
                                                                                                                                                                                                                                                                                                     22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCIKTSLTGRDKNQVEGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                   SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL
                                                                                                                                                                                                                                                                                        5 GSVVIVGRINLSGD---TAYAQOTRGEQGTOKTSHTGRDKNQVEGEVQIVSTAIOTFLAT
                                                                                                                                                                                                                                                                      3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                 DB 20; Length 665;
                                                                                                                                                                                                                                                85.8%; Score 871.5; DB 20; Length
85.2%; Pred. No. 1.8e-82;
ive 14; Mismatches 12; Indels
          ż
                                                   New hepatitis C virus covalent complexes
          Yao
                                                                     Claim 6; Page 88-90; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAY17880 standard; Protein; 216 AA
         Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      HCV NS4A-NS3 complex SEQ ID NO:4.
                                                                                                                                                                                                                                                                                                                                                                                                                  VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity 85.28
Matches 167; Conservative
         SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SCHE ) SCHERING CORP.
          Taremi
                              WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                               665 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      10-JUN-1999
         Malcolm BA,
                                                                                                                                                                                                                                                                                                                                                                          122
                                                                                                                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY17880;
                                                                                                                                                                                                                                                                                                                                                                                                                                      202
                                                                                                                                                                                                                               Sequence
                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ω
                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                          ó
```

```
NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 complex and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by MRX spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3 The covalent NS4A-NS3 complexes are more soluble, stable and active than
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVBGEVQVVSTATQSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5 GSVVIVGRINLSGD----TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                            present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV; hepatitis C virus; single chain recombinant complex; linker;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    э;
Э
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85.5%; Score 868.5; DB 20; Length 216; 85.1%; Pred. No. 7.6e-83; Live 15; Mismatches 11; Indels 3;
ż
                                                                                                                          New hepatitis C virus covalent complexes
   Yao
                                                                                                                                                                                           Claim 6; Page 76-77; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Taremi SS, Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY24946 standard; Protein; 665 AA
Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV NS4A-NS3 complex SEQ ID NO:17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182 VDFIPVESLETIMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           980S-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98WO-0524528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity 85.19
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (SCHE ) SCHERING CORP.
                                                      WPI; 1999-385385/32.
```

œ

```
WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY17884;
                                                                                                                                                                                                                                                                                                                                                                                                           182
                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAY17884
 q
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      đ
                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                           82 CVNGVCWTVYHGAGSKTLAGPKGPITOMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYL 141
                                                                                                                                                                                                                                                                                                                                                                           122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                                                                     NS4A MS3 complex comprising a central hydrophobic domain of native HCV MS4A peptide. a linker, and an HCV NS3 berine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                         81
                                                                                                                                                                                                                                                                                                   5 GSVVIVGRINLSGD -- - TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                  Gaps
                                                                            present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                 3;
                                                                                                                                                                                                                                            DB 20; Length 665;
                                                                                                                                                                                                                                           Query Match 85.5%; Score 868.5; DB 20; Length Best Local Similarity 84.7%; Pred. No. 3.7e-82; Matches 166; Conservative 15; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Yao N;
                                  New hepatitis C virus covalent complexes
                                                       Claim 6; Page 97-99; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAY24941 standard; Protein; 665 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV NS4A-NS3 complex SEQ ID NO:12
                                                                                                                                                                                                                                                                                                                                                                                                                                 VDFVPVESMETTMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                      VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (SCHE ) SCHERING CORP
          WPI; 1999-385385/32
                                                                                                                                                                                                                      665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Malcolm BA,
                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY2494]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q
οy
                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                        Š
```

```
62 SINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122 VTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                          The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   22 GSVVIVGRIILSGSGSITAYSQQTRGLIGCKITSLIGRDKNQVEGEVQVVSTATOSFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV; hepatitis C virus; single chain recombinant complex; linker;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ..
.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 20; Length 665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85.4%; Score 867.5; DB 2
84.7%; Pred. No. 4.7e-82;
tive 15; Mismatches 12
New hepatitis C virus covalent complexes
                                                               Page 85-87; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY17884 standard; Protein; 216 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex SEQ ID NO:8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     202 VDFVPVESMETIMRSP 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98US-0094331.
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-0524528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 85.4%
Best Local Similarity 84.7%
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  665 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
```

```
The present invention describes a covalent hepatitis C virus (HCV) NSAA NS3 complex comprising a central hydrophobic domain of native HCV NS4A PS3 complex comprising a central hydrophobic domain of native HCV NS3 erine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethored by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence represents a specifically claimed example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                       New hepatitis C virus covalent complexes
                                                                                Claim 6; Page 80-81; 211pp; English.
```

216 AA;

Sequence

62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL 121 5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQXTSHTGRDKNQVEGEVQIVSTATQTFLAT 61 22 GSVVIVGRIILSGSGSITAYSQQTRGLLGCKKTSLTGRDKNQVEGEVQVVSTATQSFLAT 81 3; Gaps 85.2%; Score 865.5; DB 20; Length 216; 84.6%; Pred. No. 1.6e-82; 1ive 16; Mismatches 11; Indels 3; Best Local Similarity 84.69 Matches 165; Conservative 182 Query Match 8 δ q ò 2 ò õ

Search completed: August 30, 2003, 19:12:25 Job time : 45.6227 secs

g

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 19:02:22 on:

// Search time 16.2134 Seconds
(without alignments)
1168.492 Million cell updates/sec

US-09-965-594-22 1016 1 MKKKGSVVIVGRINLSGDTA. Title: Perfect score:

......VAKAVDFIPVESLETIMRSP 197 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 seqs, 96168682 residues Searched: Total number of hits satisfying chosen parameters:

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_76:*
1: pirl:*
2: pir2:*
3: pir3:*
4: pir4:* Database

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	hypothetical prote			regulatory protein	autolysin, amidase	major acidic nucle	proteinase (EC 3.4	mosA protein - Rhi	hypothetical prote	hypothetical prote	hypothetical prote	transferrin-bindin	genome polyprotein	qenome polyprotein	genome polyprotein
VHBPLL	S57895	G90654	G85505	E64745	A11769	JC4013	PRLJHD	B53308	A81942	T08613	872606	JN0821	GNWVY	GNWVYP	GNWVNE
7	7	7	~	~	~	~	N	-	~	~	~	~	-	٦	7
301	361	452	452	452	770	1023	323	333	539	619	302	911	3411	3411	3414
7.6	7.6	7.6	7.6	7.6	7.6	7.6	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5	7.5
77.5	77.5	77.5	77.5	77.5	77.5	77	76.5	76.5	76.5	76.5	16	16	92	76	92
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

ALIGNMENTS

RESULT 1 GNWVC3 genome polyprotein - hepatitis C virus (strain HCV-1) N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstrains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstrains) Protein NSA4; nonstructural protein NSA5; nonstructural protein NS5 C; Species; hepatitis C virus C; Date: 30-Sep-1992 seequence—revision 30-Sep-1992 *text_change 19-Jan-2001 C; Accession: A39166; PQ0403; PQ040404	R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Proc. Nail. Acad. Scil. U.S.A. 88, 2451-2455, 1991 A;Title: Genetic organization and diversity of the hepatitis C virus. A;Reference number: A39166; MUID:91172826; PMID:1848704 A;Accession: A39166 A;Mocleule type: mRN A;Residues: 1-3011 cCHO>	A;Cross-references: GB:M62321; NID:g329873; PIDN:AAA45676.1; PID:g329874 R;Chan, S.W.; Mcomish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 1131-1141, 1992 A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship A;Reference number: PQ0393; MUID:g2268871; PMID:1316939 A;Recession: PQ0403 A;Molecule type: genomic RNA A;Residues: 1577-1633 cGHA>	A;Cross-references: DDBJ:D10128 A;Experimental Source: isolates E-b16 A;Accession: PQ0404 A;Status: preliminary A;Mocheule (type: genomic RNA A;Residues: 1577-1633 <ch2> A;Experimental source: isolates E-b17 C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: capsid protein; envelope protein; glycoprotein; hydrolase; nonstruc F;1-115/Product: capsid protein W *status predicted <crc> F;116-19/Product: envelope protein M *status predicted <mee> F;390-729/Product: nonstructural protein NSI *status predicted <nsi> F;730-1006/Product: nonstructural protein NSI *status Producted <nsi> F;730-1006/Product</nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></nsi></mee></crc></ch2>	F:1230-1237/Region: nuclectide-binding motif A (P-loop) F:1230-1237/Region: nuclectide-binding motif B F:1312-1317/Region: nuclectide-binding motif B F:1316-1319/Region: nuclection of F F:1616-1862/Product: nonstructural protein NS4# %status predicted <n4b> F:1863-2013/Product: nonstructural protein NS5 %status predicted <n4b> F:196.209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,207 Query Match 83.1%: Score 844.5; DB 1; Length 3011; Best Local Similarity 82.4%; Pred. No. 4.1e-68; Matches 168; Conservative 9; Mismatches 18; Indels 9; Gaps 1;</n4b></n4b>
---	---	---	---	--

---LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST 53

KKGSVVIVGRIN---

à

```
A.Note: neither anino acid nor nucleotide sequence is given
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein
C.Superfamily: hepatitis C virus genome polyprotein; glycoprotein; hydrolase; nonstruct
F.11-115/Product: capsid protein C *status predicted <CPC>
F.116-191/Product: envelope protein M *status predicted <NED>
F.116-191/Product: envelope protein M *status predicted <NED>
F.130-139/Product: nonstructural protein NSI *status predicted <NED>
F.130-129/Product: hepaclvirin *status predicted <NSI>
F.1130-1237/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b *status predicted <NNB>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NS5;
F;204-3011/Product: nonstructural protein NS5 *status predicted <NS5;
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,224C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (strain Taiwan)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
                                                                                                                                                                               genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru-
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A:Description: Genomic structure of the human prototype strain H of hepatitis C virus A:Reference number: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Wolecule type: genomic RNA
A; Residues: 1-3011 <INC>
A; Cross-references: G: MID:g329737; PIDN:AAA45534.1; PID:g329738
A; Inchauspe, G: Zebedee, S: Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus:
A; Reference number: A41546; MUID:92052256; PMID:1658800
                                                                                                                                                                                                                                                                                                    C; Species: hepatitis C virus
A; Note: host Homo sapiens (man)
C; Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C; Accession: A36814: A41546
R; Inchauspe, G; Zebedee, S; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M. submitted to GenBank, July 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
C;Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEQGIQKISHIGRDKNQVEGEVQIVSI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.
Virology 188, 102-113, 1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match 82.5%; Score 838.5; Best Local Similarity 81.4%; Pred. No. 1.4e Matches 166; Conservative 10; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A.Title: The 5'-terminal sequence of the hepatitis C virus genome.

A.Reference number: PC1284; MUID:91013116; PMID:2170712

A.Accession: PC1285

A.Molecule type: genomic RNA
A.Residues: 1-513 < CAK2>
A.Rolecule type: genomic RNA
A.Residues: 1-513 < CAK2>
A.Rolecule type: genomic RNA
A.Residues: 1-513 < CAK2>
A.Cross-references: GB:D00831; NID:9221511; PIDN:BAA00705.1; PID:9221512
A.Experimental source: isolate HC-J1
A.Experimental source: isolate HC-J1
A.Experimental source: isolate HC-J1
C.Superfamily: hepatitis C virus genome polyprotein
F.2-115/Product: capsid protein M *status predicted < CREP-
F.192-789/Product: nonstructural protein NSI *status predicted < NSI>
F.730-1006/Product: nonstructural protein NSI *status predicted < NSI>
F.1007-1615/Product: nonstructural protein NSI *status predicted < NSI>
F.1312-1317/Region: nucleotide-binding motif A (P-loop)
F.1312-1317/Region: nucleotide-binding motif B
F.1185-1317/Region: public for the follow of th
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Wicontains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu N; Ocontains: rapsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructur protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 c; species hepatitis C virus C; bate: 19-May-2000 sequence_revision 19-May-2000 *text_change 19-Jan-2001 C; Accession: S40770; PC1285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Molecule type: genomic RNA
A;Residues: 1-3011 <OKA>
A;Residues: 1-3011 <OKA>
A;Cross-references: EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221587
A;Cross-references: EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221587
B;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .005 RRGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSL.FGRDKNQVEGEVQIVST 1064
                                                                                     54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                       CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         54 AIQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83.0%; Score 843.5; DB 1; Length 3011; 82.4%; Pred. No. 5c-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Riokamoto, H. Submitted to the EMBL Data Library, March 1992 M. Reference number: $40770 A. Accession: $40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1185 CTRGVAKAVDFIPVESLETIMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       genome polyprotein - hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
                                                                                                                                                                                                                                                            114
                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                     ò
                                                                                                                                                                      g
                                                                                                                                                                                                                                                            ò
```

Length 3011; Indels

DB 1; 1.4e-67

```
80.5%;
                                                                                                                                                               76.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches 155; Conservative
                                                                                                                                                                  Best Local Similarity 76.5
Matches 156; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Accession: A38465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ဖ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    음
                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               င်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genome polyprotein - hepatitis C virus (strain JT)
N.Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4s: nonstructural protein NS4s: nonstructural protein NS5
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Accession: A45573
R; Tanaka, T; Kato, N; Nakagawa, M; Octsuyama, Y; Cho, M.J.; Nakazawa, T; Hijikata, Virus Res. 23, 39-53, 1992
A; Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: 8
A; Accession: A45573
A; Accession: Pa45573
A; Accession: Pa45573
A; Accession: Patininary
A; A; Accession: A45573
A; A; Accession: A5573
A; A; Accession: A5573
                                                     A. Accession: A4024
A. Molecule type: genomic RNA
C. Superfamily: hydrolase; nonstructural protein NSI status predicted CNEC>
F. 1907-1804 Product: nonstructural protein NSI status predicted CNSI>
F. 1007-1615/Product: nonstructural protein NSI status predicted CNSI>
F. 11307-1237/Region: nucleotide-binding motif R (P-loop)
F. 1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;204-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,233,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P-loop; polyprotein; serid
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
A; Residues: 1-3010 cTAN>
A; Residues: 1-3010 cTAN>
A; Residues: 1-3010 cTAN>
A; Cross-references: GB:DD1168; GB:DD171; NID:g221612; PIDN:BAA01943.1; PID:g221613
A; Experimental source: HCV-JT
A; Experimental source: HCV-JT
A; Note: Sequence extracted from NCB1 backbone (NCBIN:106206, NCBIP:106207)
C; Superfamily: hepatitis C virus genome polyprotein
C; Reywords: AFP; glycoprotein; Mqforlase; nucleotide binding; P-loop; polyprotein; Sc; Rej-2-185/Product: caps.id protein C status predicted <CPC>
F;16-191/Product: envelope protein M *status predicted <NES>
F;192-389/Product: nonstructural protein NSI *status predicted <NSI>
F;300-729/Product: nonstructural protein NSI *status predicted <NSI>
F;1007-1615/Product: hepacivirin *status predicted <NSS>
F;1130-1237/Region: nucleotide-binding motif B
F;1312-1317/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1065 ATOSFLATCINGVCWTVYHGAGSKTLAGPKGPITOMYTNVDQDLVGWHAPQGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1005 RRGREILLGPADSLEGRGWRLLAPITAYAQQTRGLFGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQIFLATSINGVLWIVYHGAGIRTIASPKGPVTQWYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGCPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .
6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 826.5; DB 1
Pred. No. 1.8e-66;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81.3%;
77.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 81.3
Best Local Similarity 77.9
Matches 159; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
A; Molecule type: genomic RNA
A; Residues: 1.3010 <TAX.
A; Residues: 1.3010 <TAX.
A; Cross-references: EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g329771
A; Cross-references: EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g329771
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstru
F; 2-115/Product: capsid protein C *status predicted <EP>F; 116-191/Product: major envelope protein E *status predicted <MEE>
F; 130-729/Product: nonstructural protein NS1 *status predicted <NS1>
F; 300-729/Product: nonstructural protein NS1 *status predicted <NS2>
F; 1201-137/Region: nucleotide-binding motif A (P-loop)
F; 131-1317/Region: nucleotide-binding motif A (P-loop)
F; 1316-1319/Region: nucleotide-binding motif B
F; 1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>
F; 166-1862/Product: nonstructural protein NS4b *status predicted <NAS>
F; 1014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
F; 196, 209, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, J. Virol. 65, 1105-1113, 1991
A;Title: Structure and organization of the hepatitis C virus genome isolated from h A;Reference number: A38465; MUID::91140698; PMID:1847440
A;Recession: A38465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N; Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonst protein N84s; nonstructural protein N85 (C; Species; hepatitis C virus (C; Species: 31-Mar-1992 *sequence_revision 31-Mar-1992 *text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1005 RRGKEILLGPADSLEGRGLRLLAPITAYSQQTRGLLGCIITSLTGRDKNQVEGEVQVVST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1065 ATQSFLATCVNGVCWTVYHGAGSKTLAAPKGPITQMYTNVDQDLVGWPKPPGARSLTPCT 1124
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53
                                                                                                                                                                                                                                                                                                                                        53
                                                                                                                                                                                                                                                                                                                                                                                                       RRGREILLGPADSIEGOGWRLLAPITAYAQQTRGLLGCIVTSLTGRDKNQVEGEVQVVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
F;1616-1862/Product: nonstructural protein NS4a *status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b *status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 *status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŝ
                                                                                                                                                                                                                                                        6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 80.1%; Score 813.5; DB 1; Length 3010; 76.0%; Pred. No. 2.7e-65;
                                                                                                                                                                    Score 817.5; DB 1; Length 3010;
Pred. No. 1.2e-65;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19; Indels
                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             21; Mismatches
                                                                                                                                                                                                                                               20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein - hepatitis C virus
```

Ġ,

ð 8

```
A; Neceule type: genomic RNA
A; Residues: 1-547, T', 549-621, V', 623-624, S', 626-652, 'DL', 655-761, T', 763-782 < HOW>
A; Residues: 1-547, T', 549-621, V', 623-624, S', 626-652, 'DL', 655-761, T', 763-782 < HOW>
A; Cross-references: Exbits. 61591
A; Note: this sequence is inconsistent with the nucleotide translation
A; Note: the authors translated the coden AGG for residue 43 as Pro, TGG for residue
as TTP, and TTC for residue 771 as Ser
A; Note: sequence extracted from NCBI backbone (NCBIN:121747, NCBIP:121748)
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP: 47 proprotein; hydrolase: nucleotide binding; P-loop; polyprotein; C; Status predicted < CPD>
F; 16-191/Product: envelope protein E *status predicted < NS1>
F; 192-189/Product: major envelope protein E *status predicted < NS1>
F; 130-129/Product: nonstructural protein NS1 *status predicted < NS2>
F; 130-1297/Region: nucleotide-binding motif A (P-loop)
F; 1317/Region: nucleotide-binding motif B
F; 1316-1319/Region: DEXH motif
F; 166-1862/Product: nonstructural protein NS4 *status predicted < NAA>
F; 166-1862/Product: nonstructural protein NS4 *status predicted < NAA>
F; 106-180/Product: nonstructural protein NS4 *status predicted < NAA>
F; 106-180/Product: nonstructural protein NS4 *status predicted < NAA>
F; 106-209,234,250,305,417,423,448,532,540,555,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstriprotein NS4s; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Species: J9-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predomin:
A;Reference number: JC5620; MUID:19366593; PMID:9223423
                                                                                                                                                                                                                                                                                                                                   A;Variety: isolate JX1
C;Date: 19-May-2000 #Sequence_revision 19-May-2000 #text_change 23-Mar-2001
C;Accession: 518030; S35570; A48332; S18029
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
Submitted to the EMBL Data Library, September 1991
A;Description: A whole genome of hepatitis C virus cDNA was isolated from a single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1065 ATQSFLATCVNGVCWTVYHGAGSKTLAGPRGPINQMYTNVDODLVGWQAPSGAASLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYINVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KRGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78.9%; Score 801.5; DB 1; Length 3010; Ilarity 75.5%; Pred. No. 3.3e-64; Conservative 20; Mismatches 21; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                         A; Reference number: S18028
A; Accession: S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best_Local Sim
Matches 154;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Accession: PS0086
A; Molecule type: genomic RNA
C; Comment: The cleavage sites of this polyprotein have not been determined.
C; Superfamily: hepatitis C virus genome polyprotein
F; 2-115/Product: capsid protein C *status predicted <NED>
F; 116-191/Product: major envelope protein B *status predicted <NED>
F; 190-729/Product: nonstructural protein NS2 *status predicted <NS2>
F; 1007-106/Froduct: nonstructural protein NS2 *status predicted <NS2>
F; 1230-1237/Region: nucleotide-binding motif B
F; 1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                 genome polyprotein - hepatitis C virus (strain J)
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS4s, nonstructural protein NS4b; nonstructural protein NS4b; nonstructural protein NS5b; nonstructural no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein - hepatitis C virus (isolate JK1)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4a; nonstructural protein NS4b; nonstructural protein NS5 C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A; Molecule type: genomic RNA
A; Resdues: 1:3010 < KAT7>
A; Cross references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g2216111
R; Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A; Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence varial A; Reference number: PS0085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F;1316-1319/Region: DEXH motif
F;1616-1662/Product: nonstructural protein NS4a #status predicted <N4A>
F;1663-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1005 RRGKEILLGPADSFGEQGWRLLAPITAYSQQTRGLLGCIITSLTGRDRNQVDGEVQVLST 1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRINLSGD-----TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQIFLAISINGVLWIVYHGAGIRIIASPKGPVIQMYINVDKDLVGWQAPQGSRSLIPCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1; Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          80.1%; Score 813.5; DB 1;
75.0%; Pred. No. 2.7e-65;
iive 23; Mismatches 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CTRGVAKAVDFVPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               153; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
```

113

```
F;2018-3033/Product: nonstructural protein NSS #status predicted <NO5>
F:196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,20
                                                                                                                                       66.8%;
69.3%;
                                                                                                                                                                                                                       Matches 124; Conservative
                                                                                                                                   Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNWVJ8
                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wilsons polyprotein - hepatitis C virus (isolate HC-J6)
N'Contains: capsid protein C: envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructu protein NS4 in nonstructural protein NS4 in Nov-2000
C; Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 17-Nov-2000
C; Accession: J01303
R; Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Ilzuka, H.; Machida, A.; Miyakawa, Y.
J. Gen. Virol. 72, 2697-7704, 1991
A; Title: Nucleotide sequence of the genomic RNA of hepatitis C virus sequence of the genomic RNA of hepatitis C virus genome polyprotein
A; Reference number: J01303; MUID:92044440; PMID:1658196
A; Accession: J01303
A; Molecule type: genomic RNA
A; Residues: J-3033; OKAA
A; Residues: J-3033; OKAA
A; Residues: J-3033; OKAA
A; Cossion: J01303
A; Molecule type: genomic RNA
A; Residues: J01303
A; Cossion: J01303
A; Cossion: J01303
A; Molecule type: genomic RNA
A; Residues: J01303
A; Molecule type: genomic R* Aidus predicted <CPC>
A; Experimental source: Solate Hotein R* Status predicted <NS2>
F; J011-1619/Product: monstructural protein NS4 * status predicted <NS3>
F; J011-1619/Product: nonstructural protein NS4 * status predicted <NS3>
F; J017/Product: nonstructural protein NS4 * status predicted <NS3>
F; J017/Product: nonstructural protein NS4 * status predicted <NS4>
F; J017/Product: nonstructural protein NS4 * status predicted <NS4>
F; J017/Product: nonstructural protein NS4 * status predicted <NS4>
F; J017/Product: nonstructural protein NS4 * status predicted <NS4>
F; J017/Product: nonstructural protein NS4 * status predicted <NS4>
F; J017/Product: nonstructural protein NS4 * status predicted <NS4>
F; J014-16404
                                                                                                                                                               A Experimental source: genotype 5a, which predominates in South Africa A. Experimental source: genotype 5a, which predominates in South Africa A.Note: the translation of the nucleotide sequence is not complete in this paper C. Superfamily: hepatitis C virus genome polyprotein C. Superfamily: hepatitis C virus genome polyprotein C. Superfamily: hepatitis C virus genome polyprotein C. Status predicted CEP>
F.2-115/Product: capsid protein C. *status predicted cEP>
F.116-191/Product: envelope protein E status predicted cMEE>
F.136-190/Product: major envelope protein E status predicted cMSI>
F.130-130/Product: nonstructural protein NSI *status predicted cNSI>
F.131-1007/Product: nonstructural protein NSI *status predicted cNSI>
F.131-130/Product: nonstructural protein MSI *status predicted cMA>
F.131-1318/Region: nucleotide-binding motif B
F.131-1320/Region: DEXH motif
F.1317-1320/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                73.2%; Score 743.5; DB 1; Length 3014; 68.6%; Pred. No. 6.4e-59;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1186 CTRGVAKALEFVPVENLETTMRSP 1209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 68.6
Matches 140; Conservative
                                A; Molecule type: mRNA
A; Residues: 1-3014 <CHA>
A; Cross-references: GB:Y13184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
A; Roceale type: genomic RNA
A; Residues: 2678-2754 CCHA>
A; Roclecule type: genomic RNA
A; Residues: 2678-2754 CCHA>
A; Crossurences: DDBJ:DDD1301
B; Crossurences: DDBJ:DDD1201018
A; Cross veferences: Grommun. 181, 279-285, 1991
B; Rato, N.; Octsuyama, Y.; Ohkoshi, S.; Nakazawa, T.; Mori, S.; Hijikata, M.; Shimo
Biochem. Biophys. Res. Commun. 181, 279-285, 1991
A; Title: Distribution of plural HCV types in Japan.
A; Reference number: PQ0554; MUID:92068204; PMID:1720309
A; Molecule type: mRNA
A; Residues: 2678-2729 CKAT>
A; Crossion: PQ0559
A; Molecule type: mRNA
A; Residues: 2678-2729 CKAT>
A; Crossion: PQ0550
A; Molecule type: mRNA
A; Residues: 2678-2729 CKAT>
A; Crossured: Argin genome polyprotein
C; Keywords: APP; capsid protein; envelope protein; glycoprotein; hydrolasse; nonstructorers f; 115-115/Product: envelope protein W #status predicted CMED>
F; 116-191/Product: envelope protein W #status predicted CMED>
F; 130-133/Product: monstructural protein NS1 *status predicted CNS2>
F; 1011-1619/Product: nonstructural protein MS2 *status predicted CNS2>
F; 1316-1221/Region: nucleotide-binding motif A (P-loop)
F; 1316-1221/Region: DEXH motif
F; 1657-1221/Region: DEXH motif
F; 1667-2017/Product: nonstructural protein NS4 *status predicted CNS5>
F; 108-3033/Product: nonstructural protein NS4 *status predicted CNS5>
F; 1967-2017/Product: nonstructural protein NS5 *status predicted CNS5>
F; 1967-2017/Product: nonstructural protein NS6 *status predicted CNS5>
F; 1967-2017/Product: nonstructural protein NS6 *status predicted CNS5>
F; 1967-2017/Product: nonstructural protein NS6 *status predicted CNS5>
F; 1967-2017/Product: NS60-2017/Product: NS60-2017/Product: NS6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: genomic RNA
A; Residues: 1-3033 < COKA>
A; References: GB:D10988; GB:D01221; NID:g221608; PIDN:BAA01761.1; PID:g221609
B; Chan, S; W.; McCmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap
J; Gen. Virol. 73, 1131-1141, 1992
A; Tille: Analysis of a new hepatitis C virus type and its phylogenetic relationship
A; Reference number: PQ0393; MUID:92268871; PMID:1316939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Species: hepatitis C virus
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 19-Jan-2001
C; Accession: A40250; PQ059; PQ059;
R; Okamoto, H.; Kural, K.; Okada, S.I.; Yamamoto, K.; Lizuka, H.; Tanaka, T.; Fukuda Virology 188, 331-341, 1992
A; Title: Full-length sequence of a hepatitis C virus genome having poor homology to A; Reference number: A40250; MUID:92230232; PMID:1314459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     genome polyprotein - hepatitis C virus (strain HC-J8)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonst
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                    ö
                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                79 IASPKGPVIQMYINVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                   19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP
                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                       ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 66.4%; Score 675; DB 1; Length 3033; Best Local Similarity 69.8%; Pred. No. 1.1e-52; Matches 125; Conservative 24; Mismatches 30; Indels
                                                                                                       28; Indels
Score 679; DB 1;
Pred. No. 4.7e-53;
                                                                                            27; Mismatches
```

OY 53 TATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPC 112	RESULT 14 BB1104 nitrate/nitrite sensor protein (EC 2.7.3) NWB1249 (similarity) - Neisseria meningi c;Species: Neisseria meningitidis C;Species: Neisseria meningitidis C;Species: Neisseria meningitidis C;Species: BB1104 R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B ri, H.; Oin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000 A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58. A;Reference number: A81000; MUID:20175755; PMID:10710307 A;Scatus: preliminary A;Molecule type: DNA	A; Kestudes: 1.290 vfts. A; Cross-references: GB:AE002473; GB:AE002098; NID:97226488; PIDN:AAF41629.1; PID:972 A; Experimental source: serogroup B, strain MC58 C; Genetics: A; Gene: NMB1249 C; Superfamily: nitrate/nitrite sensor protein narX C; Keywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferas: F; 395/Active site: His (phosphohistidine intermediate) *status predicted	Query Match 8.8%; Score 89.5; DB 2; Length 590; Best Local Similarity 20.8%; Pred. No. 2.9; Matches 45; Conservative 27; Mismatches 79; Indels 65; Gaps 6; Qy 28 EQGTORTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRTIASPKGPVT 87	QY 88 QMYTNVDKDLVGWQADQ	RESULT 15 C81911 nitrate/nitrite sensor protein (EC 2.7.3) NMA1418 [similarity] - Neisseria meningii C;Species: Neisseria meningitidis C;Species: Neisseria meningitidis C;Species: Neisseria meningitidis C;Date: 05-May-2000 \$sequence_revision 05-May-2000 \$text_change 02-Feb-2001 C;Accession: C81911 R;Parkhill, J; Achtman, M; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; M; Holzeyd, S.; Jagels, K; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre, Nature 404, 502-506, 2000 A;Title: Complete DNA sequence of a serogroup A strain of Neisseria menigitidis 2249; A;Reference number: A81775; MUID:20222556; PMID:10761919 A;Accession: C81911 A;Status: preliminary
OY 19 TAYAOQTRGEQGTOKTSHTGRDKNQVEGEVQ1VSTATQTFLATSINGVLWTVYHGAGTRT 78	RESULT 12 TOBB41 Delyprotein - douroucouli hepatitis GB virus A G:Species: douroucouli hepatitis GB virus A G:Species: douroucouli hepatitis GB virus A G:Species: douroucouli hepatitis GB virus A C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Nov-2000 G:Accession: ToB841 B:Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, J.K. J. Gen. Virol. 79, 41-45, 1998 A:Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A:Reference number: 216486; MUID:98120818; PMID:9460920 A:Accession: TOB841 A:Accession: TOB841 A:Molecule type: mRNA A:Residues: 1-3005 <err> A:Residues: 1-3005 <err> A:Residues: 1-3005 <err> C:Superiamily: hepatitis C virus genome polyprotein C:Superiamily: hepatitis C virus genome polyprotein</err></err></err>	Query Match 24.7%; Score 251; DB 2; Length 3005; Best Local Similarity 34.1%; Pred. No. 4.2e-14; Matches 56; Conservative 29; Mismatches 69; Indels 10; Gaps 3; Qy 33 KTSHTGRDKNQVEGEVQIVSTATOTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTN 92 III 11: II: II: II: II: II: II: II: II:	QY 93 VDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKG 152	TOBB39 TOBB39 TOBB39 TOBB39 POLYPICTEIN - marmoset hepatitis GB virus A C; Species: marmoset hepatitis GB virus A C; Species: marmoset hepatitis GB virus A C; Species: 23-301-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000 C; Accession: T08839 R; Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K. J, Gen. Virol. 79, 41-45, 1998 A; Title: Genomic analysis of two GB virus A variants isolated from captive monkeys. A; Reference number: 216486; MUID:98120818; PMID:9460920	A;Status: translated from GB/EMBL/DDBJ A;Status: translated from GB/EMBL/DDBJ A;Residues: 1-2970 < CERK> A;Cross-references: EMBL:AF023424: NID:92828597; PIDN:AAC40501.1; PID:92828598 C;Superfamily: hepatitis C virus genome polyprotein C;Superfamily: hepatitis C virus genome polyprotein C;Keywords: polyprotein Query Match Best Local Similarity 27.8%; Pred. No. 1.2e-13; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels 42; Gaps 6; Matches 62; Conservative 39; Mismatches 80; Indels

```
A:Molecule type: DNA
A;Residues: 1-590 <PAR>
A;Residues: 1-590 <PAR>
A;Residues: 1-590 <PAR>
A;Cross=references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84658.1; PID:g738007
A;Experimental source: serogroup A, strain 22491
C;Genetics:
A;Genetics: NAM4418
A;Gene: NAM419; nitrate/nitrite sensor protein narX
C;Superfamily: nitrate/nitrite sensor protein narX
C;Reywords: autophosphorylation; phosphohistidine; phosphoprotein; phosphotransferase; s
F;395/Active site: His (phosphohistidine intermediate) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                            319 DIGKYHEEIFPIEYQNETLGRLLLSFPNGISLDEDDRILLQTIGRQLGVSLAGAKQEEEK 378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            260 LLY-QTTRDLHQSYIPQQAAEHFLNRILPAVGADSGRVCLDGGSDVYVSIHHADCGTAAS 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              128 -----VIPVRRRGDSRGSLLSPRPISYLKGSSGGPLICPAGHAVGIFRAAVSTR--- 176
                                                                                                                                                                                                                                                                                                                                                             28 EQGTOKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWIVYHGAGTRTIASPKGPVT 87
                                                                                                                                                                                                                                                          Query Match 8.8%; Score 89.5; DB 2; Length 590;
Best Local Similarity 20.8%; Pred. No. 2.9;
Matches 45; Conservative 27; Mismatches 79; Indels 65; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           379 RLLAVLQERNLIAQGLHDSIAQALFFLNLQVQMLET 414
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              177
                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                               Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                   οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
```

Search completed: August 30, 2003, 19:20:32 Job time : 17.2134 secs

```
POLG_HCV1
ID POLG_HCV1
AC P26664;
                                                                                                                                                                                                                                                                                         DESTRUCTION OF STREET AND DESTREE OF STREET AND DESTREE OF STREET AND DESTREET AND 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          by genome po
h genome po
h genome po
th genome po
h genome po
h genome po
h genome po
arabidopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  hepatitis e
hepatitis e
mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               mycobacteri
neisseria m
y genome po
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 escherichia
squirrel mo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 gallus gall
mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     treponema p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   mus musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rhizobium m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         arabidopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     lactococcus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                homo sapien
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                scrofa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                musculu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        P29846 h genome p

000269 h genome p

P26661 h genome p

P2661 h genome p

P2661 h genome p

P2661 h genome a

P2661 h genome a

P2885 homo sapie

P3312 treponema

P3312 treponema

P3312 treponema

P3312 hepatitis

P3312 hepatitis

P3312 hepatitis

P3313 escherichi

P2656 lactococcu

P2656 lactococcu

P2691 escherichi

P2691 recoccu

P2691 escherichi

P2691 recoccu

P2691 escherichi

P2691 recoccu
                                                                                           August 30, 2003, 18:01:52 ; Search time 9.75674 Seconds (without alignments) 949.524 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               y genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                            1 MKKKGSVVIVGRINLSGDTA......vAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SILE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P03314
P19901
P14336
Q90627
P59240 n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Q28142
P11033
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                     Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                        127863 seqs, 47026705 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLG_HCVJ8
HHOA_ARATH
PAAD_PSEAE
DEGI_ARATH
AAMP_HUMAN
Y136_TREPA
VST2_HEVBU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ERA_MYCLE
TB11_NEIMB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_YEFV1
POLG_YEFV2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLG_TBEVW
TRY1_CHICK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLG_HCVJA
POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GRAK_MOUSE
MCP_BPF41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NX1B_BOVIN
GRAD_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VPRT_SMRVH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOSA_RHIME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WDR9_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                  Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FXH1_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NPH4_MOUSE
                                                               - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MLTD_ECOLI
                                                                                                                                                                                                                                          Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                    Minimum DB seq length: 0 Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
                                                                                                                                                     US-09-965-594-22
1016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                  SwissProt_41:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BLOSUM62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          844.5
838.5
826.5
817.5
813.5
813.5
679
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78.5
78.5
78.5
78.5
77.5
77.5
77.5
76.5
                                                                                                                                                            Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85.5
83.5
83
                                                                                                                                                                                                                          Scoring table:
                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database :
                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                          Searched:
                                                                                                Run on:
```

glibbon ape	hepatitis e	homo sapien	t qenome po	gailus gail	ustilago ma	synechocyst	triticum ae	aquifex aeo	hepatitis e	hepatitis e	oryctolagus
P21414	003200	09p253	001299	090628	P54856	001903	P46525	067531	000270	004611	P28863
POL_GALV	VST2_HEVME	VP18_HUMAN	POLG_TBEVH	TRY2_CHICK	DPM1_USTMA	SUBI_SYNY3	CS12_WHEAT	FLII_AQUAE	VST2_HEVRH	VST2_HEVMY	MM03_RABIT
٦	-	-	-	-	-	-	-	н	-	-	-
1165	629	973	3414	248	294	352	390	443	485	999	478
7.3	7.3	7.3	7.3	7.2	7.2	7.2	7.2	7.2	7.2	7.2	7.2
74.5	74	7.4	7.4	73.5	73.5	73.5	73.5	73.5	73.5	73.5	73
			_	_	33	_	_	_	_	_	

ALIGNMENTS

```
PROMITATION

DEFORMATION

DEFORMATION

DIAGRO-1992 (Rel. 23, Last sequence update)

DIAGRO-1992 (Rel. 24, Last sequence update)

DIAGRO-1992 (Rel. 1992)

DIAGRO-1993 (Rel. 1992)

DIAGRO-1993 (Rel. 1992)

DIAGRO-1993 (Rel. 1992
```

```
g
                                                                                                                              ö
                                                                                                                                                    9
                                ò
                                                     a
                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                  CAPSID PROTEIN (POTEWTIAL).
MATRIX PROTEIN (POTEWTIAL).
MAJOR ENVELORE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTEWTIAL).
RNA-DIRECTED RNA POLYMERASE (POTEWTIAL).
                                                                                                                                                                                                                                                                                               Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                             REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 844.5; DB 1; Length 3011; Pred. No. 3.4e-70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHARGE RELAY SYSTEM (BY SI
CHARGE RELAY SYSTEM (BY SI
CHARGE RELAY SYSTEM (BY SI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LINKED
                                                                              InterPro; IPR002868; HCV_NS5a.
InterPro; IPR002166; HCV_NS5a.
InterPro; IPR0010550; HG1Gase_C.
InterPro; IPR0010550; RNA_DOL_DS_PS.
InterPro; IPR007095; RNA_DOL_DS_PS.
InterPro; IPR007094; RNA_DOL_DS_PS.
InterPro; IPR0071094; RNA_DOL_PSVIT.
Pfam; PF01543; HCV_Capsid; 1.
Pfam; PF01543; HCV_Cars; 1.
Pfam; PF01560; HCV_NS1; 1.
                                                                                                                                                                                                 Pfam; PF02907; HCV_NS3; 1.
Pfam; PF02907; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS5a; 1.
Pfam; PF00151; helicase_C; 1.
Pfam; PF00998; Vital_RGRP; 1.
Probom; PD186062; HCV_NS1; 1.
                                                        HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83.1%;
82.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA; 327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     645
2041
2077
2240
2364
2789
                                                                                                                                                                                                                                                                                     SM00487; DEXDC;
                               IPR002518;
IPR004109;
IPR000745;
                                                                    IPR001490;
IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2240
2364
2789
                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                             InterPro;
InterPro;
InterPro;
                        InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                           ProDom;
SMART; S
                                                                                                                                                                                        Pfam; |
Pfam; |
                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
```

53

9; Gaps

18; Indels

Mismatches

Conservative

168;

Matches

Similarity

3 KKGSVVIVGRIN-----LSGDIAYAQQIRGEQGIQKTSHTGRDKNQVEGEVQIVST

```
54 ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-9015431: PubMed-5493270;
Kim J.L., Morganstern K.A., Gliffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonocleotide: the crystal structure provides insights into the mode
structure 6:89-100(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                         01-AUG-1992 (Rel. 23, Created)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 23, Last sequence update)
01-AUG-1992 (Rel. 42, Last annotation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NSI); Protein P1 Norther P1 Norther NS2 (P21)
(EC 3.4.21.98); Nonstructural protein NS48 (P4): Nonstructural protein
NS48 (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate H) (HCV).
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  X-FAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-97331322; PubMed-9187654;
Yao N., Hesson T., Cable M., Hoop 2., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Blol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1- FUNCTION: PROTEASE NS IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3-NS-1-FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A NS4A NS4A NS4A NS5A NS5A NS5A.

-1- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -:- FUNCTION: NSSA SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
-:- FUNCTION: NSSB IS A RNA-DEPENDENT RNA POLIMERASE THAT PLAYS AN
ESSENTIAL ROLE IN THE VIRUS REPLICATION.
-:- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or fir in P1 and Ser or Ala in P1'.
-:- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                      SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY ILIPOPROTEIN ENVELOPE, THE ENVELOPE CONSISTS OF TWO PROTEINS: ELAND E2, THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Genomic structure of the human prototype strain H of hepatitis virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
MEDLINE-92052256; Pubmed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                                                                                                                                                                         PRT; 3011 AA
                                                                                                                                                                                                                                                           174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepacivirus.
NCBI_TaxID=11108;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Prince A.M.;
                                                                                                                                                                                                                                                                                                                                                                   POLG_HCVH
ID POLG_HCVH
```

```
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1652
; 327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 838.5; DB 1;
Pred. No. 1.2e-69;
                                                 (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
               N-LINKED (

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1507
1511
1514
1514
1550
1550
1550
1584
1666
1622
1622
1635
1640
1635
1640
1635
                                                                                                                                                                                                                                                                                                                                                                                                                                                       12523335

125233335

125233335

1252323335

125232335

12523335

1253335

1253335

1253335

1253335

1253335

1253335

1253335

1253335

1253335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

125335

1253
                                          CARBOHYD
TOTURN
TUTNN
TUTNN
TUTNN
TUTNN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
STRAND
STRAND
STRAND
TURN
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
HELIX
TURN
               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Probom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Glycoprotein; Brwelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease; 3D-structure.

REMOVED FROM CAPSID PROTEIN C BY THE THILLHER I I REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
PTM: THE STRUCTURAL PROTEINS C, E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES. SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY U39. SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C.
ENVELOPE GLYCOPROTEIN E1.
ENVELOPE GLYCOPROTEIN E2.
PROTEIN P7.
NONSTRUCTURAL PROTEIN NS2.
PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS48.
NONSTRUCTURAL PROTEIN NS48.
NONSTRUCTURAL PROTEIN NS58.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. CHARGE RELAY SYSTEM (BY S. ATP (POTENTIAL).

DECH BOX.

N-LINNED (GLCNAC. . ) (PO N-LINNED (GL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IRR002521; HCV_capsid.
InterPro; IPR002521; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002531; HCV_core.
InterPro; IPR002513; HCV_NS1.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR004109; HCV_NS4.
InterPro; IPR001868; HCV_NS5a.
InterPro; IPR0070868; HCV_NS5a.
InterPro; IPR007095; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_PSvir.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01543; HCV_NS1; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01006; HCV_NS4b; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00271; helicase_C; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M67463; AAA4534.1; ...
PIR; A36814; GWWVCH.
PDB; 1HEI; 25.NOV-98.
PDB; 1A1V; 16.FEB-99.
PDB; 1A1R; 17-JUN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR001410;
IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PDB; 1A1R; 17-JUN-98
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TRANSFAC; TO4155;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         192
384
810
1052
11058
11058
11083
11165
11165
1230
1316
234
234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; InterPro; InterPro; InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CCHAIN
CCARBOHYD
C
```

```
InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                         InterPro;
InterPro;
                                                                                                                                            (nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                        ProDom; SMART; SP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                      Pfam;
                                                                                                                                                                                                                                                                                                                       Pfam;
   1064
                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMB. outstation the European Bioinformatics Institute. There are restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-slb.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                             54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                  01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 25, Last sequence update)
01-APR-1993 (Rel. 42, Last amotation update)
Genome polyprotein (Contains Capsid protein C (Core protein) (P22);
Envelope qlycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/hellcase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P60) (RNA-directed RNA pollymerase) (EC 2.7.7.48)].
Hepatitis C virus (isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN BYNELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRNA.
   9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
   Indels
   19:
                                                                                                                                                                                                                                                                                                                                                                       PRT; 3010 AA
10; Mismatches
                                                                                                                                                                                                                                                                   1185 CTRGVAKAVDFIPVENLETTMRSP 1208
                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; M84754; -; NOT_ANNOTATED_CDS
PIR; A40244; GNWVTW.
Matches 166; Conservative
                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1N64; 25-FEB-03
1NS3; 08-APR-98
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    {RNA}(N).
                                                                                                                                                                                                                                                                                                                                  RESULT 3
POLG_HCTWW
AC P29846; VTW
AC P29846; VTW
AC P29846; VTW
AC P29846; VTW
AC D1-APR-1993
DT 15-SEP-2003
DE GENOME POLYP
DE GENOME POLYP
DE (EC 34-22-
DE (EC 34-2-
DE (EC
                                                                    g
                                                                                                                                 g
                                                                                                                                                                        ò
                                                                                                                                                                                                           ద
                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                              g
                                                                                                       ŏ
```

InterPro; IPR001410; DEAD

MEROPS; S29.001; -.

```
1;
                                                                                                                                                                                                                                                                                                                                                                                                                          REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           [ POTENTIAL ]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         327047 MW; AAD267D55CDFE215 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 826.5; DB 1
Pred. No. 1.6e-68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18; Mismatches
                                                                                                                   ...uerPro; IPR007094; RNA_POL_DS_PS.
...uerPro; IPR007094; RNA_POL_DS_PS.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01549; HCV_env; 1.
Pfam; PF01560; HCV_env; 1.
Pfam; PF01560; HCV_NSI; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED (N-LINKED (N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                            PF02907; HCV_NS3; 1.
PF01006; HCV_NS4s; 1.
PF01001; HCV_NS4s; 1.
PF01506; HCV_NS5s; 1.
PF00591; helicase_C; 1.
PF00998; Viral_RGRP; 1.
                                                                                                        HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity 77.99
Matches 159; Conservative
                                                                                                                                                                                                                                                                                                                                                       SM00487; DEXDC;
IPR002521;
IPR002519;
                               IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2041 20
2077 20
2240 22
2529 25
2788 27
3010 AA;
                                                                                         IPR001490;
                                                                                                          IPR002868;
                                                                                                                                                                                                                                              PF02907; 1
                                                                                                                                                                                                                                                                                                                           Pfam; PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                   3D-structure.
INIT_MET
```

S

```
CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN (COTEWITAL).
MAJRIN PROTEIN (COTEWITAL).
MAJOR ENVELOPE PROTEIN E (POTEWITAL).
NONSTRUCTURAL PROTEIN NS.L'EZ (POTEWITAL).
PROTEASE/HELICASE NS3 (POTEWITAL).
NONSTRUCTURAL PROTEIN NS4 ROTEWITAL).
NONSTRUCTURAL PROTEIN NS4 ROTEWITAL).
NONSTRUCTURAL PROTEIN NS4 ROTEWITAL).
RNA DIRECTED RNA POLYMERASE (POTEWITAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       80.5%; Score 817.5; DB 1; Length 3010; 76.5%; Pred. No. 1.1e-67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3010 AA; 326573 MW; 94A1C77435D642BB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATP (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 1.1e
20; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED
                                                                           InterPro; IPR002518; HCV_NS2.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS3.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001409; HCV_NS4.
InterPro; IPR001608; HCV_NS5a.
InterPro; IPR001608; HCV_NS5a.
InterPro; IPR001095; RNA_POL_DSPS.
InterPro; IPR007094; RNA_POL_PSVir.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01543; HCV_capsid; 1.
Pfam; PP01539; HCV_NS1; 1.
Pfam; PP01019; HCV_NS3; 1.
Pfam; PP01001; HCV_NS3; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS3; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4; 1.
Pfam; PP01001; HCV_NS4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 76.5
Matches 156; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2529
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3D-structure,
                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
NP_BIND
SITE
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROI entry is copyright. It is produced through a collaboration
                           CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
3 KRGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST 53
                                                                                                                                                                                                                                                                                                                                                                                                                               01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SPR-1993 (Rel. 42, Last amortation update)
15-SPR-2003 (Rel. 42, Last amortation update)
16-SPR-2003 (Rel. 42, Last amortation update)
Genome polyprotein (Contains: Capsid protein C (Core protein) (P22):
Envelope glycoprotein El (GP32) (GP35): Envelope glycoprotein E2
(GP68) (GP70) (NG1): Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Heppaivirin)
(EC 3.4.22.-98): Nonstructural protein NS4A (P4): Nonstructural protein
NS4B (P27): Nonstructural protein NS5A (P56); Nonstructural protein
NS5B (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLECCAPSID COVERED BY LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLECCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                      1185 CTRGVAKAVDFVPVESMETTMRSP 1208
                                                                                                                                                                                                                                                174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A. MEDLINE-92295714; Pubmed-1318627;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D11168; BAA01943.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1A10; 25-MAR-98.
1JXP; 14-JAN-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-31642;
                                                                                                                                                                                                                                                                                                                                                                                            POLG_HCVJT
Q00269;
                                                                                 54
                                                                                                                                                                 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              <u>-</u>
                                      셤
                                                                                                                      a
                                                                                                                                                        ò
                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                              HIDDEN STANKEN STANKEN
                                                                               ò
```

ä 53

Gaps

```
1005 RRGREILLGPADSIEGQGWRLLAPITAYAQQTRGLLGCIVTSLTGRDKNQVEGEVQVVST 1064
                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                              01-40G-1992 (Rel. 23, Created)
01-40G-1992 (Rel. 23, Last sequence update)
01-5CEP-2003 (Rel. 42, Last sequence update)
65-5EP-2003 (Rel. 42, Last sequence update)
66-600 (Contains an object of the contains an object of the contains and contains and contains (GP68) (GP70) (NS1); Protein P7) (Norstructural protein NS2 (P21)
6C 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
6C 3.4.21.-); Protease/helicase NS5 (P50); Nonstructural protein
6C 3.4.21.-); Nonstructural protein NS5A (P56); 
                                                 54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                    .065 ATQSFLATCVNGVCWTVFHGAGSKTLAGPKGPITQMYTNVDQDLVGWHAPPGARSLTPCT
MEDIINE-91140699; PubMed-1847440; Takamizava A., Mori C., Fujita J., Takamizava A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J., Onlshi E., Andoh T., Yoshida I., Okayama H.; S. Andoh T., Yoshida I., Okayama H.; S. Eructure and organization of the hepatitis C virus genome isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-97015088; PubMed-8861916;
Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,
Moomaw E.W., Adachi T., Hostomska Z.;
MThe crystal structure of hepatitis C virus NS3 proteinase reveals
trypsin-like fold and a structural zinc binding site.";
Cell 87:331-342(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-96235224; PubMed-8647104;
Borowski P., Heiland M., Oehlmann K., Becker B., Korneteky L.;
"Non-structural protein 3 of hepatitis C virus inhibits
phosphorylation mediated by cAMP-dependent protein kinase.";
Eur. J. Biochem. 237:611-618(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215.
______
                                                                                                                                                                                                                                                                                                                PRT; 3010 A.A.
                                                                                                                                                                                                           1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                        174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   from human carriers.";
J. Virol. 65:1105-1113(1991).
                                                                                                                                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 1487-1500.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID-11105;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RNA](N)
                                                                                                                                                                                                                                                                                                              POLG_HCVBK
P26663;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus
                                                                                                                                                                                                                                                                               RESULT 5
                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                        g
                                                                                 a
                                                                                                                    ò
                                                                                                                                                                                        ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CAPSID PROTEIN ( POTENTIAL).
MATRIX PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 ( POTENTIAL).
PROYEASE/HELICASE NS3 ( POTENTIAL).
PROYEASE/HELICASE NS3 ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A ( POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A ( POTENTIAL).
RNA DIRECTED RNA POLYMERASE ( POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A
              LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REMOVED FROM CAPSID PROTEIN C BY
                                                                                      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        INTERPOS USY-UNITY
INTERPOS USY-UNITY
INTERPOS USY-UNITY
INTERPOS IPRO01521; HCV_capsid.
INTERPOS IPRO01521; HCV_carsid.
INTERPOS IPRO01521; HCV_core.
INTERPOS IPRO01518; HCV_NS1.
INTERPOS IPRO01518; HCV_NS1.
INTERPOS IPRO01409; HCV_NS2.
INTERPOS IPRO01409; HCV_NS4.
INTERPOS IPRO01409; HCV_NS4.
INTERPOS IPRO0166; HCV_NS4.
INTERPOS IPRO0166; HCV_NS4.
INTERPOS IPRO0166; HCV_NS5.
INTERPOS IPRO01049; RNA_POL_DS_PS.
INTERPOS IPRO01094; RNA_POL_PS_PS.
INTERPOS IPRO01094; RNA_POL_PS_PS.
INTERPOS IPRO01094; RNA_POL_PS_VIT.
Pfam; PF01543; HCV_capsid; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pfam; PP01542; HCV_COTE; 1. Pfam; PP01539; HCV_COTE; 1. Pfam; PP01539; HCV_ENY; 1. Pfam; PP01548; HCV_NS2; 1. Pfam; PP01006; HCV_NS3; 1. Pfam; PP010001; HCV_NS4s; 1. Pfam; PP01001; HCV_NS4s; 1. Pfam; PP01506; HCV_NS5a; 1. Pfam; PP019099; VITAL_RARP; 1. Pr0Dom; PD186062; HCV_NS1; 1. SNART; SMO0487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                EMBL; M58335; AAA72945.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     729
1006
1615
1615
1862
2013
369
1107
1107
11165
11165
                                                                                                                                                                                                                                                                                                                                 A38465; GNWVTC.
1A10; 25-MAR-98.
1DXP: 14-JAN-98.
1DXP: 26-APR-98.
1C2P: 15-NOV-00.
1C2P: 15-NOV-00.
1CXF: 09-APR-02.
1GX5: 09-APR-02.
1GX6: 10-APR-02.
1GW6: 10-APR-02.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1616
1863
2014
347
1083
1107
1165
1230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEROPS; S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                   PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
```

```
InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002531;
InterPro; IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR004109;
InterPro; IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P26663; LJXP.
MEROPS; S29.001; -.
MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                          PIR; A39253; GNWVCJ
HSSP; P26663; 1JXP.
                                                                                                                       Hepacivirus.
NCBI_TaxID=11116;
                                       01-AUG-1992 (
01-AUG-1992 (
28-FEB-2003 (
                           POLG_HCVJA
P26662;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
               RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                ATQTFLATSINGVLWTVYHGAGTRIIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                              CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                          53
                                                                                                                                                                                                                                                                                                                                                                                         3 KKGSVVIVGRIN------1,SCDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                           (POTENTIAL)
                                                                        (POTENTIAL).
(POTENTIAL).
                                                                                           (POTENTIAL)
                                                                                                               (POTENTIAL)
                            POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                             6
                    POTENTIAL)
                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                              80.1%; Score 813.5; DB 1; Length 3010; 76.0%; Pred. No. 2.7e-67;
                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                  327189 MW; F8422D5ECCFDFD9C CRC64;
                                (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                               (GLCNAC.
(GLCNAC.
(GLCNAC.
                    (GLCNAC.
                                                           (GLCNAC.
                                                                                           (GLCNAC.
                                                                         (GLCNAC
                                                                                (GLCNAC
                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                            21; Mismatches
      N-LINKED
                                                                                                              N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                           645
2041
2047
2240
2240
2288
2788
1047
1063
1074
                                                                                                                                                                         1688
                                                                                                                                                                                                                                                                                                                                                  3010 AA;
                                                                                                                                                                                                                                                                                                                                                                     Local Similarity
mes 155; Conserv
                                                                                                                                                                                                                11101
11104
11120
11122
11129
11135
                                                                                                                                                                                                                                                                          1158
1162
1165
1168
                                                                                                                                                                  1075
1077
1082
1086
1090
1093
                                CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                           CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                   54
                                                                                                                                                                                                                                                                                                                                                                                                                                              114
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1125
                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                          CARBOHYD
                                                                                                                                             STRAND
STRAND
STRAND
                                                                                                                                                                                                         STRAND
STRAND
TURN
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                       TURN
STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                              STRAND
TURN
STRAND
                                                                                                                                                                  TURN
                                                                                                                                  STRAND
                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                                         STRAND
                                                                                                                                                                                                                                                                                                                                            STRAND
                                                                                                                                                                                HELTX
                                                                                                                                                                                                                                                                                                                                                                       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                  IURN
                                                                                                                                                                                       TURN
                                                                                                                                                                                                    TURN
g
                                                                                                                                                                                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   셤
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                      01-AUG-1992 (Rel. 23, Last Sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
Genome polyprotein [Contains: Capsid protein C (Core protein [P.2]);
Envelope glycoprotein El (GP32) (GP35); Envelope glycoprotein E2
(GP69) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS5A (P56); Nonstructural protein NS5B (P66); Nonstructural protein NS5B (P66); CP7); Nonstructural protein NS5B (P66); Nonstructural protein NS5B (P66); Nonstructural protein NS5B (P66) (P70) (RM2-directed RNA Polymerase) (EC 2.7.7.48)).
NSSB (P66) (P70) (RM2-directed RNA Polymerase) (EC 2.7.7.48)).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Obsobil S., Shimotohno K.;

"Molecular structure of the Japanese hepatitis C viral genome.";

"Molecular structure of the Japanese hepatitis C viral genome.";

"FEBS LECT. 2806.1325.328 [1991].

"INSTAND NOS MALL PROPEINS NS2A, NS2B, NS4A AND NS4B ARE
HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION.

NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.

"CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.

"CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BY A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROTEIN ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
MEDLINE-91088550; PubMed-2175903;
Kato N., Hijikata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
Japanese patients with non-A, non-B hepatitis.;
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISCUSSION OF SEQUENCE.
MEDLINE-91192160; PubMed-1849488;
Kato N., Hijikata M., Nakagawa M., Ootsuyama Y., Muraiso K.,
3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
                                       (Rel. 23, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; D90208; BAA14233.1; -.
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002166;
```

```
ATQTFLATSINGVLWTVYHGAGIRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3 KKGSVVIVGRINLSGD-----TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST 53
                                                                                                                               9; Gaps
                                                                                                                                                                                                                                            PROTEASE/HELICASE NS3 (POTENTIAL)
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL)
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL)
RNA-DIRECTED RNA POLYMERASE (POTENTIAL)
                                                                                                                                                                                                                                                                                 POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MW; AA993794F46DB185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           80.1%; Score 813.5; DB 1;
75.0%; Pred. No. 2.7e-67;
Live 23; Mismatches 19;
                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                  (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (GLCNAC
                                                                                                                                                                                                                                                                                                                    ATP (POTENTIAL)
DECH BOX.
                                                                                                                                                                                                                                                                                                                                               N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_core; 1.
Pfam; PF01539; HCV_NS3; 1.
Pfam; PF01001; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4; 1.
Pfam; PF01006; HCV_NS4b; 1.
Pfam; PF01005; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF0100998; Viral_RAP; 1.
Pfam; PF00998; Viral_RAP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 153; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2240
2788
3010
                                                                                                                                                                                                                                                                                                                                                                                                                                                           645
2041
2077
                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1125
                                                                                                                                                                                                                                                                                          ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                            ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                   FRANSMEM
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                CHAIN
                                                                                                                                                                                                                                                                          CHAIN
   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE-9204440; PubMed-1658196;
A Okamoto H., Okada S.-I., Sugiyama Y., Kurai K., Lizuka H.,
A Machida A., Miyakawa Y., Mayumi M.,
I mucleotide sequence of the genomic RNA of hepatitis C virus isolated
T rom a human carrier: comparison with reported isolates for conserved
and divergent regions. "12:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 72:8697-2704(1991).
J. Gen. Virol. 73:8697-2704(1991).
J. Gen. Virol. 73:8697-2704(1991).
C -!- FUNCTION: PROFIEING NESA, MSZB, NSZB, RELATED FUNCTION.
NS3 AND NS5 AMY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 AMY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
NS3 AND NS5 AMY PLAY A ROLE IN THE VIRAL RNA REPLICATION.
OF CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral prostition, Cys or Thr in Pl and Ser or Ala in Pl'.
C CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
                                                                                                                                                                                                        01-MCG-1992 (Rel. 23, Created)
01-MCG-1992 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation Edge of protein (Rel. 41, Rel. 41, Rel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            {RNA}(N).
-1-SUBONT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPPOTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                            PRT; 3033 A.A.
1185 CTRGVAKAVDFIPVESMETTMRSP 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001650; Helicase_C.
InterPro; IPR007099; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
Pfam; PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid.
HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV_env.
HCV_NS1.
HCV_NS2.
HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_RdRP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV_NS4a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL, D00944, BAA00792.1; -. PIR, JQ1303, JQ1303. HSSP; P27958; IHEI.
                                                                                                                                                                            STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002518;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; IPR001410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEROPS; S29.001; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11113;
                                                                                                                                                                         POLG_HCVJ6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEROPS;
                                                                                                                                                POLC_HCVJ6
                                                                                                                 RESULT 7
```

```
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1094 LAGSRGPVTQMYSSAEGDLVGWPSPPGTKSLEPCTCGAVDLYLVTRNADVIPARRRGDKR 1153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               79 IASPKGPVTQMYTNVDKDI.VGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                        CABSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROFEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                       Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; APP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
INIT_MET 1 CELLULAR AMINOPEPTIDASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0; Caps
                                                                                                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                  .) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 1; Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28; Indels
                                                                                                                                                                                                                                                                    POTENTIAL.
CHARGE RELAY SYSTEM (BY S
CHARGE RELAY SYSTEM (BY S
CHARGE RELAY SYSTEM (BY S
ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F957F5C1A273BE9E
                                                                                                                                                                                                                                                                                                                                                                                              (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                       (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     No. 9e-55;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                27; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 679;
                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                                                             N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        329165 MM;
                                              F01006; HCV_NS4a; 1.
F01001; HCV_NS4b; 1.
F01506; HCV_NS5a; 1.
F00271; hclicase_C; 1.
PF00998; Viral_RGRP; 1.
PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           66.8%;
69.3%;
                   HCV_NS1; 1.
HCV_NS2; 1.
HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 69.34
Matches 124; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              649
1091
2038
2811
                                                                                                                                                                                                         733
1010
1619
1866
2017
3033
369
1087
1111
1111
1169
1124
11323
                                                                                                    ProDom; PD186062; HCV_I
SMART; SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ¥
                                                                                                                                                                                                                                                                             1087
1111
1111
1234
1320
196
209
234
305
417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3033
                                                 Pfam; PF01006; F
Pfam; PF01001; F
Pfam; PF01506; F
                                         PF02907;
                                                                                            Pfam; PF00998
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                    TRANSMEM
ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                   ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
POLG_HCVJ8
                                                                                                                                                                                             CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                 CHAIN
                                                                                                                                                                                                                                                            CHAIN
                                                                                 Pfam;
                                                                                                                                                                             CHAIN
                                                                                                                                                                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Š
```

3033 AA

STANDARD;

POLG_HCVJ8 P26661;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL cutstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-cib.ch).
         01-703 (Rel. 23, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
28-FEB-2003 (Rel. 41, Last annotation update)
66-000 polyprotein [Contains: Capsid protein C (Core protein E)
67-00 (Rel.); Protein P7; Nonstructural protein NS2 (P21)
6EC 3-4.22-7); Protease/helicase NS3 (P70) (Hepacivirin)
6EC 3-4.21-98); Nonstructural protein NS4A (P4); Nonstructural protein
NS4B (P27); Nonstructural protein NS5A (P55); Nonstructural protein
NS5B (P65) (P70) (RN-4directed RNA pollymerase) (EC 2.7.7.48)].
Hepatliis C virus (1solate HG-38) (HCV).
                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE-92230223; PubMed-1314459;
Okamoto H., Kurai K., Okada S.-I., Yamamoto K., Lizuka H., Tanaka T.,
Fukuda S., Tsuda F., Mishiro S.;
"Full-length sequence of a hepatitis C virus genome having poor
homology to reported isolates: comparative study of four distinct
genotypes.":
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the viral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENTELOPE. CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MENA.
                                                                                                                                                                                                                                                                                                                                                                                                          -i- FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRARE.FELATED FUNCTION.
NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RMA REPLICATION.
-i- CATALYTIC ACTIVITY: Hydrolysis of four peptide bonds in the viral precursor polyprotein, commonly with Asp or Glu in the P6 position, Cys or Thr in P1 and Ser or Ala in P1.
-i- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -1 - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                InterPro; IPR002166; HCV_RGRP.
InterPro; IPR007095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSvir.
Pfam; PF01543; HCV_Cap91d; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR000745; HCV_NS4a.
InterPro; IPR001490; HCV_NS4a.
InterPro; IPR001690; HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_capsid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV_env
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; D10988; BAA01761.1; -. PIR; A40250; GNWVJ8.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Viral RdRP;
                                                                                                                                                                                                                                                                                                                                                                                             Virology 188:331-341(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_core;
HCV_env; 1
HCV_NS1; 1
HCV_NS2; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV_NS5a;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410;
InterPro; IPR002522;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro; IPR002521;
InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               THEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEROPS; S29.001; -.
                                                                                                                                                                                                                 Hepacivirus.
NCBI_TaxID=11115;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PF01506;
PF00998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    140250; G
P27958;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA (N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; A40250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
```

```
.;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19 TAYAQOTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRSP 197
                                                                                       CELLULAR AMINOPEPTIONSE.
CAPSID PROTEIN ( POTENTIAL).
MATHER PROTEIN ( POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Arabidopsis thaliana (Mouse-ear cress).
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0; Gaps
     SMART; SM0487; DEDC; 1.
SMART; SM0487; DEDC; 1.
SPOIYPFOTEIN; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.
                                                                                                                                                                                                                                                  POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL) (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 3033;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             30; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC. . ) (POT W; 1A173E7E3381FD1A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HHOA_ARATH
095EL7; 049507;
10-607-2001 (Rel. 40, Created)
116-007-2001 (Rel. 40, Last sequence update)
116-007-2001 (Rel. 40, Last sequence update)
116-007-2001 (Rel. 40, Last annotation update)
116-007-2001 (Rel. 40, Last precursor (EC 3.4.21.-).
116-007-2001 (Rel. 40, CRES)

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             66.4%; Score 675; DB 1;
69.8%; Pred. No. 2.1e-54;
tive 24; Mismatches 30
                                                                                                                                                                                                                                                                                                                                                                                             (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GICNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                  N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED
                                                                                                                                                                                                                                                                                                                                                             N-LINKED
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                330177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 69.8% les 125; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                  BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 9
HHOA_ARATH
                                                                                                                          CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                       CHAIN
                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Aatches
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
```

```
WEBLINGTON ON WEILINGTON ON WEBLIAND ON WEILINGTON ON WEILINGTON ON WEBLIAND THE WEBLIAND THE WEBLIAND ON WEBLIAND
```

```
SEQUENCE FROM N.A.
STRAIN-cv. Columbia;
MEDLINE-20363099; PubMed-10907853;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 51; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Res. 7:217-221(2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE OF 104-118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         190 LET 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       eurosids II; Bra.
NCBL_TaxID-3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDTTDDT
       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                        entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11;
                                           Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KISPSVVYIEAIELPKISSGDILTDEENGKIEGTGSGFVWDKLGHI-----VTNYHVIA 148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         52 STATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTP 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              202 VVLGTSNDLRVGQSCFAI----GNPYGYENTLTIGVVSGLGREIPSPNGKSISEAIQTD 256
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2 KKKGSVVIVGRINL----SGDTAYAQQTRGEQGT-----QKTSHTGRDKNQVEGEVQIV 51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
STRAIN-ATCC 15692 / PAO1;
STRAIN-ATCC 15692 / PAO1;
STRAIN-LOUGH STRAIN STRAIN STRAIN-ATCC 15692 / PAO1;
STOWN STRAIN-STRAIN STRAIN STRAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               112 CTCGSSDLYLVTRHADVIPVRRRGDSRG------SLLSPRPISYLK----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hydrolase; Serine protease; Chloroplast; Thylakoid; Transit peptide.
TRANSIT 1 26 CHLOROPLAST (POTENTIAL).
non-profit institutions as long as its content is in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----GSSGGPLLCPAGHAVGIFRAAVSTR--GVAKAVDF-IPVESLETIM 194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                257 ADINSGNSGGPLLDSYGHTIGVNTAIFTRKGSGMSSGVNFAIPIDTVVRTV 307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Nature 406:959-964(2000).
-1- SIMILARITY: BELONGS TO THE POLYPRENYL P-HYDROXYBENZOATE / PHENYLACRYLIC ACID DECARBOXYLASES FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1; Length 321;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   89; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      R -> G (IN REF. 1).
68DB81E0BD27A7A7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Probable aromatic acid decarboxylase (EC 4.1.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No. 0.43;
                                    modified and this statement is not removed.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        THYLAKOID.
PROTEASE HHOA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               31; Mismatches
                                                                                                                                                                     EMBL; AF114386; AAF24060.1; -...
EMBL; AL021710; CAA16717.1; ALT_SEQ.
EMBL; AL161548; CAB78839.1; ALT_SEQ.
MEROPS; S01.279; -..
InterPro; IPR001940; Protease2C,
InterPro; IPR001254; Ser_protease_Try.
Pfam; PF00089; trypsin; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8.9%; Score 90;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLY-GLU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (Rel. 40, Created)
(Rel. 40, Last sequ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pseudomonadaceae: Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34691 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pseudomonas aeruginosa
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          321 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID-287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-0CT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAAD_PSEAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACT_SITE
ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SECUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAAD_PSEAE
10 PAAD_R
10 PAAD_R
10 COHNO
DT 16-CCT
DT 16-CCT
DE Probab
GN PA4019
GN PA4019
GN PA4019
GN PREDLIN
RR SEQUEN
RR GATCH
RC GATC
   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
```

```
between the Swiss Institute of Bioinformatics and the EMEL outstainon the European Bioinformatics and the EMEL outstainon the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        92 NVDKDLVGWQAPQGSRSLTP----CTCGSSDL-----YLVTRHADVIPVRRGDS 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         81 QND-----WMAPPASGSSAPNAMVICPCSTGTLSAVATGACNNLIERAADVALKER---- 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          138 RGSLLSPR--PIS-----YLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVES 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  29 QEEREVHFLISKAAQLVMATETDVALPAKPQAMQAFLTEYCGAAAGQI-----RVFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEGL_ARATH STANDARD; PRT; 437 AA.
DEGL_ARATH STANDARD; PRT; 437 AA.
D22609; 09LR85;
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 41, Last annotation update)
Protease Do-like 1, chloroplast precursor (EC 3.4.21.-).
DEGPL OR DEGP OR A73G27925 OR K16N12.18.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; eudicotyledons; Core eudicots; Rosidae; eurosids II: Erassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               43 QVEGEVQ-IVSTATOTFLATSINGVL-------WIVYHGAGTRTIASPRGPVTQMYT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Kaneko T., Katoh T., Sato S., Nakamura A., Asamizu E., Tabata S.;
"Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN-cv. Columbia;
Kisselbach T., Bystedt M., Schroeder W.P.;
Submitted (JUL-2000) to the SWISS-PROT data bank.
-1- FUNCTION: SERINE PROTEASE THAT IS REQUIRED AT HIGH TEMPERATURE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A., AND CHARACTERIZATION.
MEDLINE-98175982; PubMed-9507020;
Itzhaki H., Naveh L., Lindahl M., Cook M., Adam Z.;
Identification and characterization of DegP, a serine protease associated with the luminal side of the thylakoid membrane.";
J. Biol. Chem. 273:7094-7098(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               55;
                                                                                                                                                                                                                                                                                                             EMBL; AE004818; AAG07406.1; -.
PIR; H83144; H83144.
InterPro; IRN003382; Flavoprotein.
Pfam; PF0244; Flavoprotein; 1.
Hypothetical protein; Lyase; Decarboxylase; Complete proteome.
SEQUENCE 209 AA; 22367 MM; 01FD081CC495D3F6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 209;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          61; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0.67;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          16; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Match 8.4%; Score 85.5; Local Similarity 27.9%; Pred. No. 0
```

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
      SO THE THE THE TRANSPORT OF THE PRESENCE OF THE PROPERTY OF TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMED cutstation the European Bloinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ------0A 102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     103 PQGSRSLTPCTCGSSDLYLV------TRHADVIPVRRRGDSRGSLLSPRPI 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            150 VPQGSGSGFVWDKQGHIVTNYHVIRGASDLRVTLADQTTFDAKVVGFDQDKDVAVLRIDA 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              210 PK--NKLRPIPVGVSADLLVGQKVFAIGNPFGLDHTLTTGVISGLRREIS--SAATGRPI 265
                   CAN DEGRADE BETA-CASEIN.
ENZYME REGULATION: INHIBITED BY PHENYLMETHYLSULFONYL FLUORIDE AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Transit peptide; Chloroplast; Thylakoid. CHLOROPLAST (POTENTIAL).
MAY BE INVOLVED IN THE DEGRADATION OF DAMAGED PROTEINS. IN VIVO,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -> SSTLFLHSPPSSHL (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SYL-----KGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDF-IPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            266 QDVIQTDAAINPGNSGGPLLDSSGTLIGINTAIXSPSGASSGVGFSIPVDIV 317
                                                                               -1 - SUBCELLULAR LOCATION: BOUND TO LUMINAL SIDE OF THE THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 52;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
CHARGE RELAY SYSTEM (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  8.2%; Score 83.5; DB 1; Length 437; ilarity 26.2%; Pred. No. 2.5; Conservative 17; Mismatches 55; Indels 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      70 VYHGAGIRIIASPKGPVTQMY-------TNVDKDLVGW-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1497BlAB3F5FF2A4 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2).
V -> I (IN REF. 2).
V -> S (IN REF. 2).
G -> R (IN REF. 2).
G -> D (IN REF. 2).
LL -> HF (IN REF. 2).
L -> V (IN REF. 2).
I -> V (IN REF. 2).
I -> V (IN REF. 2).
O -> E (IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROTEASE DO-LIKE 1. SERINE PROTEASE.
                                                                                                                      -1- INDUCTION: By heat shock.
-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C
-1- SIMILARITY: Contains 1 PDZ/DHR domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HSPPSSQLSNST
                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro; 1901.179;
InterPro; 1PR001478; PDZ,
InterPro; 1PR001254; Ser_protease_C.
InterPro; 1PR001254; Ser_protease_Try.
Pfam; PP00595; PDZ; 1.
Pfam; PP00899; trypsin; 1.
PRINTS; PR00834; PROTEASES_C.
SMART; SM00228; PDZ; 1.
PROSITE; PS50106; PDZ; 1.
Hydrolase; Serine protease; Transit pept
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            THYLAKOID
                                                                                                                                                                                                                                                                                                                                                                    EMBL; AF028842; AAC39436.1; -.
EMBL; AP000371; BAB02539.1; -.
EMBL; AP001302; BAB02539.1; JOINED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             46213 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          36
50
60
64
69
355
381
                                                              O-PHENANTHROLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Matches 45; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                   MEROPS; S01.279;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hydrolase;
TRANSIT
TRANSIT
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACT_SITE
ACT_SITE
ACT_SITE
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
CONFLICT
CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAMP_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 12
셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMED cutstation the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      68 WIVYHGAGTRIIASPKGPVTQMYINVDKDLVGWQAPQGSRSL----TPCTCGSSDLYLV 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   197 WMEWH------PRAPVILAGT-ADGNTWMWKVPNGDCKTFQGPNCPATCGR----- 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   123 TRHADVIPVRRR---GDSRGS-----1LSPRPISYLKGSSG--GPLLCPA------ 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   241 ----VLPDGKRAVVGYEDGTIRIWDLKQGSPIHVLKGTEGHQGPLTCVAANQDGSLILT 295
                                                                                                                                                                                                                                     Gaps
                                                                         Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ------GHAVGIFR----AAVSTRGVAKAVDFIPVESL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 296 GSVDCQAKLVSATTGKVVGVFRPETVASQPSLGEGEESESNSVESL 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HEPARIN-BINDING (POTENTIAL).
POLY-GLU.
WD 1.
WD 2.
WD 3.
WD 4.
WD 5.
WD 6.
WD 7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8.2%; Score 83; DB 1; Length 452; 25.3%; Pred. No. 2.9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DA1413D25EB236C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PIR; 139383; 139383.

Genew; HGNC:18; AAMP.
MIM; 603488; -.
GC; GC)008201; F:heparin binding activity; TAS.
Interpro; IPR001680; WD40.
Pfam; PF00400; WD40; 8.
SMART; SM00320; WD40; 8.
15-JUL-1998 (Rel. 36, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Angio-associated migratory cell protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13; Mismatches
                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE; PSO0676; WD_REPEATS_1; 1.
PROSITE; PSSO082; WD_REPEATS_2; 6.
PROSITE; PSSO294; WD_REPEATS_REGION; 1.
Repeat; WD_repeat.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           49015 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; M95627; AAA68889.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           452 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
hes 42; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
```

452 AA

PRT;

AANP_HUMAN STANDARD; 013685; 15-JUL-1998 (Rel. 36, Created)

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
    STATE THE PROPERTY OF STATE OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 165 CGGGGGSSSSSSCIHIWLVPGGTGNNGNCGCGGGGGSSSSSSCIHIKVENTDEQFL 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  109 GIGVYKHCVNGAGSSSTGTTASPSTETCSQHAT----LVGGTSKPFWLVPGGTGNNGNCG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 C----GSSDLYLVIRHADVIP-----VRRRGDSRGSLLSPRPISYLK----- 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           63
                                                                                                                                                                                                                                                                                                                      Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S., Wenter J.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SGDTAYA-----QQTRGEQGTQKTSH----TGRDKNQVEGEVQIVSTATQTFLATSI-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      54 AGSKLYATNGRLWEKELNGTGSWQKVSSSSVPTDSDK----KVMSIATDGNTFVLACVP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 -NGVLWTVYHGAG---TRIIASPKGPVTQMYINVDKDLVG-----WQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             71; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Science 281:375-388(1998).
-!- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
                                                                                                                                                                                      Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HYPOTHETICAL LIPOPROTEIN TP0136. Na ACYL DIGLYCERIDE (POTENTIAL). GLY/SER RICH. GLY/SER-RICH. GLY/SER-RICH. POLY-SER. POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       *Complete genome sequence of Treponema pallidum, the syphilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (Potential).
-1. SIMILARITY: BELONGS TO THE TP013X FAMILY OF LIPOPROTEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.7%; Score 78.5; DB 1; Length 485; 23.4%; Pred. No. 8.3; 1ve 16; Mismatches 77; Indels 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  C7A4CEEDC7DC5CED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AE001199; AAC65137.1; ALT_INIT.
TIGR; TP0136; -.
Hypothetical protein; Lipoprotein; Membrane; Signal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DMGEGYVVTTKHLYTKNGSSSAGPAQCPGGGGG 258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----GSSGGPLLCPAGHAVG 167
                                           16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hypothetical lipoprotein TP0136 precursor.
485 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            or send an email to license@isb-sib.ch)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   POTENTIAL
                                                                                                                                                                                                                                                                                                     MEDLINE-98332770; PubMed-9665876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    48984 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23
485
24
178
210
267
327
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          444
485 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
Matches 50; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Complete proteome.
                                                                                                                                                              Freponema pallidum
                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                         NCBI_TaxID-160;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 spirochete.
  Y136 TREPA
                                                                                                                                                                                      Bacteria;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LIPID
DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
                                                                                                                                         TP0136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CHAIN
  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to licenselisb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94 DKDLVGWQAPQGSRSLTPCTCGSSDLYLV···TRHADVIPVRRGDSRG-SLLSPRPISY 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  417 ENA----QQDKGIAIPHDIDLGESRVVIQDYDNQHEQDRPTPSPAPSRPFSVLRANDVLW 472
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10 VGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQV----EGEVQIVSTATQTFLA----T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -----RIIASPKG-PVTQMYTNV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                        THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
                                                                                                                                                                                                                                                                                   MEDLINE-92024067; PubMed-1926770;
Tan A.W., Smith M.M., Guerra M.E., Huang C.-C., Bradley D.W.,
Fry K.E., Reyes G.R.;
"Hepatitis E virus (HEV): molecular cloning and sequencing of the (HI)-length viral genome.";
Virology 185:120-131(1991).
-i- FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  150 LK------GSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPV 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          7.7%; Score 78.5; DB 1; Length 660;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               473 LSLTAAEYDQSTYGSSTGPVY - - VSDSVTLVNVATGAQAVARSLDWTKV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          85; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POTENTIAL.
STRUCTURAL PROTEIN 2.
5832A013CCC4A61C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-DEC-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last sequence update)
01-DEC-1994 (Rel. 28, Last sequence update)
Structural protein 2 precursor (ORF2).
Hepatitis E virus (strain Burma) (HEV).
Viruses: ssraw positive-strand viruses, no DNA stage;
Hepatitis E-like viruses.
$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          42; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 12;
099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      62 SINGV-----LWTVYHGAGT------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          660 AA; 70978 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M73218; AAA45736.1; -.
PIR; C40778; VHWWH2.
InterPro; IPR004261; SP2.
Pfam; PF03014; SP2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          45; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                               NCBI_TaxID=31767;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VST2_HEVPA
P33426;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VST2_HEVPA

ID VST2_HE

DT PST2_HE

DT 01-FEB

DT 01-FEB

DE Structt

OS Hepatii

OC Hepatii

OX NCBI_TT
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its map by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94 DKULVGWQAPQCSRSLTPCTCGSSDLYLV---TRHADVIPVRRRGDSRG-SILLSPRPISY 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10 VGRINLSGDTAYAQQTRGEQGTQKTSHTGRDKNQV----EGEVQIVSTATQTFLA----T 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62 SINGV-----LWTVYHGAGT------93
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57; Gaps
                                                 Table S.A., Emerson S.U., Reyes G.R., Tsareva T.S., Legters L.J., Malik I.A., Iqbal M., Purcell R.H.;

"Characterization of a prototype strain of hepatitis E virus.";

Proc. Natl. Acad. Sci. U.S.A. 89:559-563(1992).

-I. FUNCTION: CONTAINS A HIGH BASIC AMINO ACID CONTENT SUGGESTING THAT IT MAY BE INVOLVED IN THE ENCAPSIDATION OF THE GENOMIC RNA BY EFFECTIVELY NEUTRALIZING THE NEGATIVELY CHARGED RNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         150 LK------GSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPV 187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LSLTAAEYDQSTYGSSTGPVY -- VSDSVTLVNVATGAQAVARSLDWTKV 519
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match

7.7%; Score 78.5; DB 1; Length 660;
Best Local Similarity 19.7%; Pred. No. 12;
Matches 45; Conservative 42; Mismatches 85; Indels 5
                                                                                                                                                                                                                                                                                                                                                                                                                       BY SIMILARITY.
STRUCTURAL PROTEIN 2.
8085BC53CFB46FD3 CRC64;
                                    MEDLINE-92115700; PubMed-1731327;
                                                                                                                                                                                                                                                                                                                                                                                                                       1 22 B1
23 660 S7
660 AA; 70980 MW;
                                                                                                                                                                                                                                                                                                                                              EMBL; M80581; AAA45727.1; -.
Interpro; IPR004261; SP2.
Pfam; PF03014; SP2; 1.
[1]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                      Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
```

Search completed: August 30, 2003, 19:13:50 Job time: 10.7567 secs

```
Q91fh8 mucosal dis
086579 hepatitis c
081756 hepatitis c
Q91fe5 hepatitis c
Q96188 hepatitis c
003463 hepatitis c
09608 hepatitis c
Q91008 hepatitis c
Q91078 hepatitis c
Q91078 hepatitis c
Q91077 hepatitis c
Q91078 hepatitis c
Q91078 hepatitis c
Q91078 hepatitis c
                                                                             (without alignments)
1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                   1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197
                                                                  August 30, 2003, 19:00:22 ; Search time 37.5921 Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Description
                                                                                                                                                                                                                  830525
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                             Total number of hits satisfying chosen parameters:
                                                                                                                                                                                          830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SUMMARIES
                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                             OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               036579
081756
091FE5
09ELS8
003463
036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9PWU9
Q91RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q91RT5
Q91RR5
Q91RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    091RT9
091RR3
091RR4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0915н8
                                                                                                                                                         BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                          sp_archea:*
sp_bacteria:*
sp_fung1:*
sp_human:*
sp_invertebrate:*
sp_manmal:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_vertebrate:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_rvirus:*
sp_bacteriap:*
sp_archeap:*
                                                                                                                                                                                                                                                                                                                                                                                                       sp_organelle:*
sp_phage:*
                                                                                                                                                                                                                                     Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                               sp_plant:*
sp_rodent:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                       sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length DB
                                                                                                                                                                                                                                                                                                                 SPTREMBL_23:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2436
3011
3011
3011
3011
3015
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
181
181
181
181
181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    862.5
8844.5
8844.5
8844.5
8843.5
8841.5
881.5
839
                                                                                                              Title:
Perfect score:
                                                                                                                                                          Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               837
837
837
                                                                                                                                                                                                                                                                                                                 Database :
                                                                                                                                     Sequence:
                                                                                                                                                                                           Searched:
                                                                  Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Result
```

•	O91rs1 hepatitis c		Q91rt1 hepatitis c	-	_	hepatitis						hepatitis					hepatitis			-				hepatitis	hepatitis	hepatitis	hepatitis	hepatiti	hepatitis
	Q91RS1	Q91RQ8	Q91RT1	Q91RR0	036609	Q91RR6	Q91RS9	Q91RS3	09DIT6	091RT4	091RS8	Q91RT3	Q91RS5	Q91RS7	Q91RT0	091RS2	Q91RS6	Q9QP61	091RS4	068533	Q91RR7	Q91RT6	036610	091RT8	070818	070817	Q91RR9	Q9DTE2	Q99AU2
			181 12																										
1	ς.	ď.	82.3	ä	ä	ä	'n	ä	ä	ä	81.9	Ξ.	;	;		Η.	;	;	Ξ.	H	4	ä	;	;	;	٦.	ä	81.3	٦.
	836	836	836	30	835.5	834	834	œ	832.5	832	832	832	832	832	832	832	ω	830,5	ထ	829.5	829	829	829	œ	827.5	2	æ	826.5	56
,	17	18	19	20	21	22	23	24	25	26	27	28	53	30	31	32	33	34	35	36	37	38	39	7	41	42	43	44	45

ALIGNMENTS

```
RESULT 1

O91FH8 PRELIMINARY; PRT; 4040 AA.

C 091FH8 PRELIMINARY; PRT; 4040 AA.

C 091FH8 PRELIMINARY; PRT; 4040 AA.

DT 01-CCT-2000 (TrEMBLE-1.15, Last sequence update)

DT 01-CCT-2000 (TrEMBLE-1.15, Last sequence update)

DT 01-MAC-2003 (TrEMBLE-1.15, Last sequence update)

DE Genome polyprotedin.

S Mucosal disease virus.

OC VILUSES: SRRAM POSITIVE-strand viruses, no DNA stage; Flaviviridae;

OC VILUSES: SRRAM POSITIVE-strand viruses, no DNA stage; Flaviviridae;

OC PESTIVIUS.

NCB1_TAXID-11099;

RN BEDINE-202324848; PubMed-10864644;

RA Lav V.C., Zhong W. Skelton A. Ingravallo P., Vassilev V.,

RA Lav V.C., Zhong W. Skelton A., Ingravallo P., Vassilev V.,

RA Lav V.C., Zhong W. Skelton A., Ingravallo P., Vassilev V.,

RA Donis R.O., Rong Z., Lau J.Y.;

RY Gependent bovine virual diarrhea virus.*;

LA VIROL 74:6339-6347(2000).

RR Gependent bovine virual diarrhea virus.*;

LA VIROL 74:6339-6347(2000).

RR Lai V.C. H., Hong Z.;

RR SEQUENCE FROM N.A.

RR Lai V.C. H., Hong Z.;

RR SEQUENCE FROM N.A.

RR Lai V.C. H., Hong Z.;

RR SEQUENCE FROM N.A.

RR SEQUENCE FROM N.A.

RR SEQUENCE FROM N.A.

RR HORD S. S110013.-

DR RICEPPO: IPR001410; DESP.

DR INTERPRO: IPR001405 HVZ-NS3.

DR INTERPRO: IPR00105; Myb_DNA_Dinding.

BR InterPro: IPR00105; Myb_DNA_Dinding.

BR InterPro: IPR00105; Myb_DNA_Dinding.

BR InterPro: IPR00105; Wyb_DNA_Dinding.

BR InterPro: IPR001095; RNA_POL_DS_PS.

BR InterPro: IPR001095; RNA_POL_DS_PS.

BR InterPro: IPR001095; Wiral_RRP; 1.

BR Ffam; PF00998; Viral_RRP; 1.
```

```
Query Match
  Pfam;
                       Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 3
Q81756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ద
                                                                                                                                                                                                                                                                    ï
                                                                                                                                                                                                                                                                                                                                                                                                                62 SINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYL, 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          122 VTRHADVIPVRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKA 181
                                                                                                                                                                                                                                                                                                                                   10 GSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSLTGRDKNQVEGEVQIVSTATQTFLAT 69
                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LTABLE THE VICTOR IN THE CONTRACT OF THE NUCLEOCAPSID COVERED BY A LIPOPROTEIN HAND GLYCOPROTEIN E. THE VICTOR THE VICTOR IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL: AF009606; AAB66324.1; --
HSSP: P27958; 1HEI.
                                                                                                                                                                                                                                                                                                             5 GSVVIVGRINLSGD---TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S.M.
               SMART: SM00487; DEXDC: 1.

PROSITE; P800037; MTE_1: 1.

PROSITE; P550037; MTE_1: 1.

PROSITE; P550521; RDRP_POSITIVE: 1.

PROSITE; P550521; RDRP_VIRAL: 1.

PROSITE; P500531; RNASE_12_2: 1.

ATP-Inding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerse; Transferase.

SEQUENCE 4040 AA: 453073 MM; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                    3; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kolykhalov A.A., Agapov E.V., Blight K.J., Mihalik K., Feinstone Rice C.M.;
                                                                                                                                                                                                                         Query Match 84.9%; Score 862.5; DB 12; Length 4040; Best Local Similarity 88.7%; Pred. No. 1.9e-72; Matches 173; Conservative 5; Mismatches 14; Indels 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "Transmission of hepatitis C by intrahepatic inoculation with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAN-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Helicase_C.
RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV_capsid.
HCV_core.
HCV_env.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-97373636; PubMed-9228008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4b.
HCV_NS5a.
HCV_RdRP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR000745; HCV_NS4a. IPR001490; HCV_NS4b.
  PRINTS; PR00729; CDVENDOPTASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182 VDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD.
InterPro; IPR002522; HCV_C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR007095; I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 IPR001650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. STRAIN-H77;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       036579
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DESTRUCTION OF STREET OF S
    a
                                                                                                                                                                                                                                                                                                                                                                                                                           ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
Ĥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23
Prime Pro1551; MCV_COV. 1.

R Prime Pro1559; MCV_COV. 1.

R Prime Pro1569; MCV_COV. 1.

R Prime Pro1569; MCV_COV. 1.

R Prime Pro1569; MCV_COV. 1.

R Prime Pro1000; MCV_COV. 1.

R Prime Pro1000; MCV_COV. 1.

R Prime Pro1009; MCV_COV. 1.

R Pro099; Viral_EMRP; 1.

R Pro099; Viral_EMRP; 1.

R Pro099; Viral_EMRP; 1.

R PRO517E; PS5052; RDRP_POSITIVE; 1.

R ROSITE; PS5052; RDRP_POSITIVE; 1.

R PRO517E; PS5052; RDRP_VIRAL; 1.

R RATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; W RATP-binding; Coat protein; Postofin; Postofin; Coat protein; Mainterase; Transferase; Transf
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
Choo O.-L., Richman K., Han J.;
Choo O.-L., Richman K., Han J.;
The nucleotide sequence of the Hepatitis C viral genome.";
Submitted (MAY-1990) to the EMBL/GenBank/DDBJ databases.
EMBL, M32084; AAA45677.1; -.
HSSP; P27958; 1A1V.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83.5%; Score 848.5; DB 12; Length
82.4%; Pred. No. 2.8e-71;
ive 10; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q81756 PRELIMINARY; PRT; 2436 AA. Q81756; OBLIS6; OBLIVOV-1996 (TrEMBLrel. 01, Created) OI-NOV-1996 (TrEMBLrel. 01, Last sequence update) OI-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410; DEAD.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR000445; HCV_NS4a.
InterPro; IPR001490; HCV_NS4a.
InterPro; IPR001490; HCV_NS5a.
InterPro; IPR001669; HCV_NS5a.
InterPro; IPR001669; HCV_NSFa.
InterPro; IPR001095; RNA_POI_DS_PS.
InterPro; IPR007095; RNA_POI_DS_PS.
InterPro; IPR007095; RNA_POI_DS_PS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Genome polyprotein (Fragment).
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity 82.48 Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01560; HCV_NS1; 1
Pfam; PF01538; HCV_NS2; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepacivirus.
NCBI_TaxID-11103;
```

```
09ELS8
09ELS8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 5
09ELS8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              sequence.";
J. Gen. Virol. 82:1291-1297(2001).

I. Gen. Virol. 82:1291-1297(2001).

I. SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID IS A COMPLEX OF PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA (BY SIMILARITY).

EMBL, AF211632, AAF81759.1;

HSSP, P27958; LAV.

InterPro; IPR001415; CytC_heme_bind.

InterPro; IPR001415; DabD.

InterPro; IPR001522; HCV_capsid.

InterPro; IPR002521; HCV_capsid.

InterPro; IPR002521; HCV_capsid.
                                                                                                                                                                           SMART; SMO487; DEXDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50507; RDRP_VIRAL; 1.
APP-binding; Coat protein; Envelope protein; Glycoprotein; Helicasc; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA Polymerase; Transferase; Transmembrane.
NON_IER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. MEDLINE-21262212; PubMed-11369872; Lanford R.E., Lee H., Chavez D., Guerra B., Brasky K.M.; "Infectious cDNA clone of the hepatitis C virus genotype 1 prototype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83.1%; Score 844.5; DB 12; Length 2436;
82.4%; Pred. No. 5.1e-71;
tive 9; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2436 AA; 264734 MW; D7B9872900BE3125 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-0CT-2000 (TrEMBLrel. 15, Created)
01-0CT-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       735 CTRGVAKAVDFIPVENLETTMRSP 758
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   174 STRGVAKAVDFIPVESLETTMRSP 197
Pfam; PF01005; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF00271; Helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                             2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR000745;
InterPro; IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                             2436
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9IFE5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT A CONTROL OF THE CONTROL OF T
DR WEEN THE STATE OF THE STATE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

```
54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SMART; SM00487; DEXDC. 1.
PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
ATP-Dinding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA-directed RNA polymerase; Transmembrane.
SEQUENCE 3011 AA; 327124 MW; 2489CE74AC864E58 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
Genome polyprotein.
18-partial C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepatius.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 83.1%; Score 844.5; DB 12; Length 3011; Best Local Similarity 82.4%; Pred. No. 6.7e-71; Matches 168; Conservative 9; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR001266; HCV_RGRP.
InterPro; IPR001659; Hellcase_C.
InterPro; IPR001095; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_DS_PS.
InterPro; IPR007094; RNA_POL_PSvir.
Pfam; PF01543; HCV_capsid; 1.
Pfam; PF01559; HCV_cnc; 1.
Pfam; PF01550; HCV_NS; 1.
Pfam; PF01005; HCV_NS; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS5a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01001; HCV_NS4a; 1.
Pfam; PF01006; HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-colonel;
```

nterPro;

```
"Nucleotide sequences of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions.";
J. Gen. Virol. 72:2697-2704(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE-92044440; PubMed-1658196;
Okamoto H., Okada S., Sugiyama Y., Kurai K., Iizuka H., Machida A.,
Miyakawa Y., Mayumi M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                        Okamoto H., Kanai N., Mishiro S.;
"Full-length nucleotide sequence of a Japanese hepatitis C virus isolate (HC-JI) with high homology to USA isolates.";
Nucleic Acids Res. 20:6410-6410(1992).
"The 5'-terminal sequence of the hepatitis C virus genome.";
Jpn. J. Exp. Med. 60:167-177(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROSITE: PSS0507; RDRP POSITIVE: 1.
PROSITE; PSS0521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein;
Hydrolase; Monstructural protein; Polyprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE-94174722; PubMed-7510436;
Mink M., Benichou S., Madaule P., Tiollais P., Prince A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hydrolase; Nonstructural process; restractions RNA-directed RNA polymerase; Transferase; Transmembrane enninence 3011 AA; 327112 MW; 97E9052C0250463B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_NS1.
InterPro; IPR004109; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001490; HCV_NS4.
InterPro; IPR001566; HCV_NS4.
InterPro; IPR001566; HCV_RAP.
InterPro; IPR001509; HAD_DS_PS.
InterPro; IPR001509; NA_PO1_DS_PS.
InterPro; IPR007094; RNA_PO1_DS_PS.
InterPro; IPR07094; RNA_PO1_DS_PS.
InterPro; IPR07194; HCV_COTe; I.
Pfam; PF01543; HCV_COTe; I.
Pfam; PF01543; HCV_COTe; I.
Pfam; PF01559; HCV_NS1; I.
Pfam; PF01559; HCV_NS2; I.
                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE-93117120; PubMed-1335573;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pfam; PF01006; HCV_NS43; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01506; HCV_NS54; 1.
Pfam; PF000271; helicase_C; 1.
Pfam; PF00998; Viral_RGRP; 1.
ProDom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         interPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SM00487; DEXDC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; PF01538; H
Pfam; PF02907; H
Pfam; PF01006; H
Pfam; PF01001; H
Pfam; PF01506; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-HC-J1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Inchauspe G.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-HC-J]
                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-HC-J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Okamoto H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART;
              SOW WENT THE TRANSPORT OF THE TRANSPORT 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      protein; Glycoprotein; Helicase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 12; Length 3011;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Okamoto H., Okada S., Sugiyama Y., Yotsumoto S., Tanaka T.,
Yoshizawa H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.1%; Score 844.5; DB 12; Length
81.9%; Pred. No. 6.7e-71;
.ive 11; Mismatches 17; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE; PS00190; CYTCCHROME_C; 1.
PROSITE: PS50507; RDRP_POSITIVE; 1.
ATP-binding; Coat protein: Envelope protein; Glycoprotein Hydrolase; Monstructural protein; Polybrotein: RNA-directed RNA Polymerase; Transferase; Transmembrane. SEQUENCE 3011 AA; 327107 MW; AGBECFSA1B3EE13F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TIEMBLEEL. 01, Created)
01-NOV-1996 (TIEMBLEEL. 01, Last sequence update)
01-NAR-2003 (TIEMBLEEL. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT; 3011 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro; IPR002668; HCV_NSSA.
InterPro; IPR001066; HCV_NSSA.
InterPro; IPR001050; Halicase_C.
InterPro; IPR001095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_DS_PS_INC_PFam; PF01543; HCV_core; I.
Pfam; PF01543; HCV_core; I.
Pfam; PF01548; HCV_env; I.
Pfam; PF01589; HCV_NSI; I.
Pfam; PF01589; HCV_NSI; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRGVAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-91013116; PubMed-2170712;
                                                                                                                   IPR002518; HCV_NS2.
IPR004109; HCV_NS3.
IPR000745; HCV_NS4a.
IPR001490; HCV_NS4b.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF02907; HCV_NS3; 1.
Pfam; PF02090; HCV_NS4s; 1.
Pfam; PF01001; HCV_NS4s; 1.
Pfam; PF01506; HCV_NS5s; 1.
Pfam; PF001506; HCV_NS5s; 1.
Pfam; PF001998; Viral_RGRP; 1.
Probom; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity 81.9
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                          PR002521;
                                                                      PR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepaclvirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-HC-J]
                                                                                                                                                   InterPro;
InterPro;
                                                                                             InterPro;
InterPro;
                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  174
```

Query Match

ò g ò g à g õ Q03463;

RESULT (
003463

AC 0003463

AC 0000

DT 011

DT 011

DT 011

CO VIII

CO Hele

CO H

Helicase;

S

```
Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (nterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             interPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Q9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      O9PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09PWX5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The state of the s
DR KEE DR R SO W KEE
                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                  1005 RKGREILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                               54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT 113
                                                                                                                                                                                                                                                              114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                   53
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-97385173; PubMed-9238047; Yanaqi M., Purcell R., Emerson S.U., Bukh J.; Yanaqi M., Purcell R., Emerson S.U., Bukh J.; "Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a
                                                                                                   3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            chimparzee.;
Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).
Proc. Natl. Acad. Sci. U.S.A. 94:8738-8743(1997).
LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GIYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
PROTEIN C AND MRNA (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Genome polyprotein.
Hepatitis C virus strain H77.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                     DB 12; Length 3011;
                                                           6.
                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-JAN-1998 (TrEMBLrel. 05, Created)
01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                         19;
                Score 843.5; DB 13
Pred. No. 8.4e-71;
8; Mismatches 19.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 3011 A.A.
                                                                                                                                                                                                                                                                                                                                                             Interpro; IPR001210; BRD.
Interpro; IPR001221; HCV_capsid.
Interpro; IPR002521; HCV_capsid.
Interpro; IPR002519; HCV_core.
Interpro; IPR002519; HCV_core.
Interpro; IPR002519; HCV_core.
Interpro; IPR002519; HCV_NS1.
Interpro; IPR001409; HCV_NS4.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001409; HCV_NS4b.
Interpro; IPR001409; HCV_RGRP.
Interpro; IPR001569; HCV_RGRP.
Interpro; IPR001569; HCV_RGRP.
Interpro; IPR0017095; RNA_pol_DS_PS.
Interpro; IPR001044; RNA_pol_DS_PS.
Interpro; IPR001044; RNA_pol_DS_PS.
Interpro; IPR001044; RNA_pol_DS_PS.
Interpro; IPR001044; RNA_pol_PSVIT.
Pfam: PF01543; HCV_capsid; 1.
                                                                                                                                                                                                                                                                                                                                          174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AF011751; AAB67036.1; -. HSSP; P27958; 1HEI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV_core; 1.
HCV_env; 1.
HCV_NS1; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV_NS5a; 1.
                83.08;
                                        82.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS2; 1.
HCV_NS3; 1.
HCV_NS4a; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV_NS4b;
                                        Best_Local Similarity 82.4
Matches 168; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepacivirus.
NCBI_TaxID-63746;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PF01506;
PF00271;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PF01542;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PF01539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF01560;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PF01538;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF01006;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-H77
                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
036608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                   යි
                                                                                                                                                                                    õ
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                            ò
```

```
Ä
                                                                                                                                                                                                                                                                                                                         1005 RRGOEILLGPADGMVSKGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST 1064
                                                                                                                                                                                                                                                                                       53
                                                                                                                                                                                                                                                                                                                                                                    54 ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                    114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV
                                                                                                                                                                                                                                                                                     3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHIGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                            9; Gaps
                                                       PROSITE; PSSOSO7: RDRP_POSITIVE; 1.
PROSITE; PSSOSO1: RDRP_VIRAL: 1.
ATP-binding; Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3011 AA; 327112 MW; 0B75E6B81CB5C198 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-99420396; PubMed-10489358;
Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.;
"Hepatitis C virus: an infectious molecular clone of a second major genotype (2a) and lack of viability of intertypic la and 2a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Viruses; saRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                       DB 12; Length 3011;
                                                                                                                                                                                                 Ouery Match 82.8%; Score 841.5; DB 12; Length Best Local Similarity 81.9%; Pred. No. 1.3e-70; Matches 167; Conservative 10; Mismatches 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT; 3015 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequ
01-MAR-2003 (TrEMBLrel. 23, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV_NS3.
HCV_NS4a.
HCV_NS4a.
HCV_NS4b.
HCV_NS5a.
HCV_NS5a.
Pfam; PF00998; Viral_RdRP; 1.
ProDom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDC; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Virology 262:250-263(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR002521;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IPR002519;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002518;
IPR004109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        IPR002166;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID-11103;
```

```
Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           091RR8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               091RRB
                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
Q91RR8
   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1129 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1188
                                                                                                                                                                                                                                                                                          ij
                                                                                                                                                                                                                                                                                                                                                       54, ATQTFLATSINGVLWTVYHGAGTRTIASPKGPVTQMYTNVDKDLVGWQAPQGSRSI,TPCT 113
                                                                                                                                                                                                                                                                                                                                                                                                114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                               53
                                                                                                                                                                                                                                                                                                             3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
--- SUBUNIT: THE VIRIN OF THIS VIROS IS A NOCLEGOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                                                                                                                                                                                                                                         9; Gaps
                                                                                                                                                                               PROSITE; PSO0392; DDC_GAD_HDC_YDC; 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
ATP-binding; Coat protein; Brotein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase; Transmembrane.
SEQUENCE 3015 AA; 328159 MM; B7D23BC1F190663A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. SEQUENCE FROM N.A. MEDLINE-99420396; Pubmed-10489358; Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.; Yanaqi M., Purcell R.H., Emerson S.U., Bukh J.; "Hepatitis C virus: an infectious molecular clone of a second major "Hepatitis C virus: an infectious molecular clone of a second major genocype (2a) and lack of viability of intertypic la and 2a Limeraq.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 12; Length 3015;
                                                                                                                                                                                                                                                                   82.8%; Score 841.5; DB 12; Length
81.9%; Pred. No. 1.3e-70;
Live 10; Mismatches 18; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 3015 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                    174 STRGVAKAVDFIPVESLETTMRSP 197
           InterPro; IPR002129; Pyridoxal_dec.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR007094; RNA_pol_PSvir.
InterPro; IPR001650; Helicase_C
                                       Fram; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01542; HCV_capsid; 1. Pfam; PF01550; HCV_NS1; 1. Pfam; PF01506; HCV_NS2; 1. Pfam; PF01006; HCV_NS4; 1. Pfam; PF01001; HCV_NS4; 1. Pfam; PF01506; HCV_NS5a; 1. Pfam; PF00998; Viral_KdRP; 1. Pr009m; PD0989; Viral_KdRP; 1. Pr000m; PD186062; HCV_NS1; 1. Pr000m; PD186062; HCV_NS1; 1.
                                                                                                                                                                                                                                                                  Ouery Match 82.81
Best Local Similarity 81.91
Matches 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9PWU9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             60M460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                             ð
                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
```

```
1129 GGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGLFRAAV 1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   114 CGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAV 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1009 RRGQEILLGPADGMVSKGWRLLAPITAYAQOTRGLLGCIITSLTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KKGSVVIVGRIN-----LSGDTAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSITE; PS00392; DDC_GAD_HDC_YDC; 1, PROSITE; PS50507; RDRP_POSITIVE; 1, PROSITE; PS50521; RDRP_VIRAL; 1. ATP-binding; Coat protein; Envelope protein; Glycoprotein; Hellcase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Frament).
Hepatitis C virus.
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           82.8%; Score 841.5; DB 12; Length 3015;
81.9%; Pred. No. 1.3e-70;
ive 10; Mismatches 18; Indels 9;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                328084 MW; E309F6318067D6CD CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hydrolase; Nonstructural protein; Polyprotein;
RNA directed RNA polymerase; Transferase; Transmembrane
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1189 CTRGVARAVDFIPVENLGTTMRSP 1212
                                                                                                                                                                                                                                                                                                                                    InterPro: IPRO01203; PAILGASE_C.
InterPro: IPRO01203; PAILGASE_C.
InterPro: IPRO01203; PAILGASE_C.
InterPro: IPRO01203; RNA_pol_DS_PS.
InterPro: IPRO01095; RNA_pol_DS_PS.
InterPro: IPRO01095; RNA_pol_DS_PS.
InterPro: IPRO01504; RNA_pol_DS_VIT.
Pfam; PF01560; HCV_core; 1.
Pfam; PF01560; HCV_NS1; 1.
Pfam; PF01065; HCV_NS3; 1.
Pfam; PF01006; HCV_NS3; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS4b; 1.
Pfam; PF01001; HCV_NS5a; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              174 STRGVAKAVDFIPVESLETTMRSP 197
                                                                                        HCV_capsid.
HCV_core.
HCV_env.
HCV_NSI.
HCV_NSI.
HCV_NSI.
HCV_NSI.
HCV_NSI.
HCV_NSI.
HCV_NSI.
EMBL; AF177039; AAF01181.1; -. EMBL; AF177037; AAF01179.1; -.
                                                                   IPR001410; DEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 81.9:
ses 167; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEXDc;
                                                                                        IPR002522; H
IPR002521; H
IPR002519; H
                                                                                                                                                       IPR002531; H
IPR002518; H
                                                                                                                                                                                                                                                                          IPR002868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3015 AA;
                                             HSSP; P2795B; 1HEI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SM00487;
                                                                                                              InterPro;
InterPro;
                                                                                                                                                                                                                                                  InterPro;
InterPro;
InterPro;
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                interPro;
                                                                                                                                                                                                                            InterPro;
                                                                                           InterPro;
                                                                                                                                                            InterPro;
                                                                                                                                                                                 InterPro;
                                                                                                                                                                                                         InterPro;
```

113

Ĥ

~

```
79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLIPCTCGSSDLYLVTRHADVIPVRRKGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4 TAYAQOTRGLLGCIXTSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRT 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          124 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFKAAVCTRGVAKAVDFIPVENLETTMRS 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-Pt.47;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AR169241; AAK54566.1;
InterPro; IPR004109; HCV.NS3.
Pfam; PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AFS69238, AAK4563.1;
InterPro, IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
19123 MW; 1CAE817345ED809D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      181 181 181 18084 MW; 3B5E8161F2100A72 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   091RR2;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment).
                                                                                                                                                                                                                                      Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 837; DB 12;
Pred. No. 8.5e-72;
2; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AA.
                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              82.48;
92.18;
                                                                                                                                                                         091RR5;
01-DEC-2001 (TrEMBLrel. 19,
01-DEC-2001 (TrEMBLrel. 19,
01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    164; Conservative
                                                                                                                                                                                                                                                                                             NS3 protease (Fragment). Hepatitis C virus.
                                                                                                                                                    PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181
181 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-Pt. 3U;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                 091RR5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q91RR2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
Q91RR2
                                                                                       RESULT
Q91RR5
                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCOCC DETAILS ACCOCC OS SELECTION OF SELECT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79 IASPKGPVTOMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79 IASPKGPVTQMYTNVDKDLVGWQAPQGSRSLTPCTCGSSDLYLVTRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4 TAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVSTAAQTFLATCINGVCWTVYHGAGTRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19 TAYAQQTRGEQGTQKISHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GSLLSPRPISYLKGSSGCPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWIVYHGAGTRT
                                                                        "Genetic Diversity and response to IFN of the NS3 Protease Gene from Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF864205; ARK646601; -
InterPro; IPR004109; HCV_NS3.
Pfam, PF02907; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Holland-Staley C.A., Kovari L.C., Golenberg E., Mayers D.L.;
"Genetic Diversity and response to IrN of the NS3 Protease Gene from
"Clinical Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL. ARF86218, ARK5443.1: -.
InterPro: IPR004109; HCV_NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  82.6%; Score 839; DB 12; Length 181; 92.7%; Pred. No. 5.5e-72; ive 1; Mismatches 12; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 181;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12; Indels
                                                                                                                                                                                                                                                                                                                                                   181 181
181 AA: 19130 MW: 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181 AA; 19130 MW; 85D91869299B7C35 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 839; DB 12;
Pred. No. 5.5e-72;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                     82.68;
92.78;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-DEC-2001 (TrEMBLrel. 19, 01-DEC-2001 (TrEMBLrel. 19, 01-MAR-2003 (TrEMBLrel. 23, NS3 protease (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches 165; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-Pt. 4;
                                                                                                                                                                                                                                                                                                                     NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Protease.
                                                                                                                                                                                                                                                                                                Protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      091RT5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SO ON SET THE SET THE SO ON SET THE SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

Gaps

```
NCBI_TaxID-11103;
                                    STRAIN-Pt.4B
                                                                                                                                  Protease.
                                                                                                                                          NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                         Ouery Match
                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                            Search con
Job time
  g
                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                     g
                                   ċ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64 IASPKGPVIQMYINVDKDLVGWPAPQGARSLIPCTCGSSDLYLVTRHADVIPVRRRGDSR 123
                                                                                                      79 IASPKGPVIQMYINVDKDLVGWQAPQGSRSLIPCICGSSDLYLVIRHADVIPVRRRGDSR 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4 TAYAQQTRGLLGCIITSLTGRDKNQVBGEVQIVSTATQTFLATCINGVCWTVYHGAGTRT 63
                                                          19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT 78
                                                                       4 TAYAQQIRGLLGCIXISLIGRDKNQVEGEVQIVSTAXQIFLATCINGVCWTVYHGAGIRI 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               139 GSLLSPRPISYLKGSSGGPLLCPAGHAVGIFRAAVSTRGVAKAVDFIPVESLETTMRS 196
                                                                                                                                                    19 TAYAQQTRGEQGTQKTSHTGRDKNQVEGEVQIVSTATQTFLATSINGVLWTVYHGAGTRT
                                                                                                                                                                                                                                                                                                                                                                             STRAIN-Pt.1C;
Holland-Staley C.A., Kovarl L.C., Golenberg E., Mayers D.L.;
Holland-Staley C.A., Kovarl L.C., Golenberg E., Mayers D.L.;
Holland-Strains of the Hepatitis C Virus.";
Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
EMBL, AF369214; AR454539.1;
InterPro; IPR004109; HCV_NS3.
Pfam; PF02907; HCV_NS3; 1.
                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0; Gaps
                                                                                                                                                                                                                                                                                        NG3 protease (Fragment).

Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NS3 protease (Fragment).
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                    .,
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 837; DB 12; Length 181;
Pred. No. 8.5e-72;
2; Mismatches 12; Indels
           Score 837; DB 12; Length 181; Pred. No. 8.5e-72;
                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181 181
181 AA; 19114 MW; BEIDOB542F014E86 CRC64;
                                                                                                                                                                                                                                                 091RT9;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
NS3 protease (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Created)
Last sequence update)
Last annotation update)
82.4%; Scur.
92.1%; Pred. No. o..
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 181 AA
                                                                                                                                                                                                                                       181 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match 82.4%;
Best Local Similarity 92.1%;
Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-DEC-2001 (TrEMBLrel. 19,
01-DEC-2001 (TrEMBLrel. 19,
01-MAR-2003 (TrEMBLrel. 23,
                                   Matches 164; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                       PRELIMINARY;
                       Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                091RR3
                                                                                                                                                                                                                                       Q91RT9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             091RR3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                           g
                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                            g
                                                                                                       õ
                                                                                                                                                                           셤
                                                                                                                                                      ò
```

Search completed: August 30, 2003, 19:18:21 Job time : 37.5921 secs

```
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ALIGNMENTS
                                                                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145273
AR145271
AR145255
AR125259
AR127810
BD081911
AR127809
BD081910
AR110828
BD081910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AF009606
                      em_htg_inv:*
em_htg_other:*
em_htg_mus:*
                                                                                                                                     em_htgo_other:*
                                                                                                             em_htgo_hum:*
em_htgo_mus:*
                                                       em_htg_pln:*
em_htg_rod:*
em_htg_mam:*
                                                                                         em_htg_vrt:*
           em_htg_hum:*
                                                                                                                                                                                                                          Query
Match Length DB
em_vi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   864.5
864.5
863.5
863.5
862.5
861.5
860.5
860.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            857.5
857.5
856.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    853.5
853.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 1
                                                                                                                                                                                                                                                                                                        Result
                                                                                                                                                                                                                                                                                                                                                          Command line parameters:
-MODEL-frame+_p2n_model -DEV-x1p
-Q-/cgn2_L/USPTO_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-Q-/cgn2_L/USPTO_spool/US09965594/runat_29082003_151919_28310/app_query.fasta_1.2872
-Q-/cgn2_L/USPTO_spool/US09965594/runat_29082003_15.00PcL-0 -LOOPEXT-0
-UNITS-b1ts -START-1 -END--1 -MATRIX-b10sum62 -TRANS-human40.cdi -LIST-45
-DOCALIGN-200 -THR_SCORE-PCT -THR_ANX-100 -ALIGN-15 -MODE-LOCAL
-UNFNT-Ptc -NORM-ext -HBAPSIZE-S00 -MILEN-0 -MXZLEN-200000000
-USER-USO9965594_GCR.1 1 14686_@tunat_2908003_151919_28310 -NCPU-6 -ICPU-3
-NO_MARP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-EGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELDEXT-7
                                                                                         (without alignments) 3147.423 Million cell updates/sec
                                                                                                                                  1016
1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETTMRSP 197
                                                                            August 30, 2003, 19:18:33; Search time 2560.57 Seconds
         GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                      - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                     2888711 seqs, 20454813386 residues
                                                                                                                                                                                                                                                                                                                Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                              0.5
0.5
0.5
                                                                                                                                                                              , Ygapext
, Ygapext
, Fgapext
, Delext
                                                                                                                                                                                                                                                                                 Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                         US-09-965-594-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9b_ba:*
9b_htg:*
9b_on:*
9b_on:*
9b_pat:*
9b_ph:*
9b_ph:*
9b_ph:*
9b_pr:*
9b_pr:*
9b_vi:*
9b_un:*
9b_vi:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_ov:*
em_pat:*
em_ph:*
em_pl:*
em_ro:*
em_ro:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             em_hum: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          em_in:*
em_mu:*
                                                                                                                                                                  BLOSUM62
Xgapop 10.0 ,
Ygapop 10.0 ,
Fgapop 6.0 ,
Delop 6.0 ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                em_om:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            em_or:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GenEmbl:*
                                                                                                                                    Perfect score:
                                                                                                                                                                     Scoring table:
                                                       OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Database :
                                                                                                                                                 Sequence:
                                                                                                                                                                                                                                      Searched:
                                                                             Run on:
```

AR145253 Sequence AR145266 Sequence AR145261 Sequence AR145269 Sequence AR268278 Pestiviru AR145257 Sequence

Sequence Sequence Sequence Sequence

Sequence

Sequence

Sequence

Sequence

Sequence

Description

Sequence

AR145260 Sequence AR14521 Sequence AR145273 Sequence AR145273 Sequence AR145255 Sequence AR15280 Sequence BD01911 Hepatitis AR127809 Sequence BD01911 Hepatitis AR127809 Sequence BD01911 Hepatitis AR127809 Sequence

BD069982 Functiona AF009606 Hepatitis AR110831 Sequence

BD069985 Functiona

AR118686 Sequence

ARI18692 Sequence 106440 Sequence 54 109329 Sequence 10 ARI18696 Sequence

109331 Sequence 15 M32084 Hepatitis (AR118703 Sequence

```
Unknown.
Unclassified.
 AR145264
AR145264.1
                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                               ď
                             Unknown
                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCUS
DEFINITION
ACCESSION
VERSION
                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                   Pred. No.:
                                     ORGANISM
                                                                 AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 3
AR145268
 ACCESSION
                                                                                           JOURNAL
          VERSION
KEYWORDS
                                                         REFERENCE
                                                                                                    FEATURES
                              SOURCE
                                                                                                                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                       ద
                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                             ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PAT 08-AUG-2001
           PAT 20-APR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              892
                                                                                                                                                                                                                                                                                                                                                                                                                  ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                               81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              833 CITICGCCCCGGCCCATIICCTACCTAAAAGCTCCTCGGGGGGTCCGCTGTIGTGCCCC
                                                                                                                                                                                                                                                                    GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                           AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                               SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                 1 (bases 1 to 12734)
Hong,Z., Lai,W.C.H. and Lau,J.Y.N.
Hepatitis C virus protease-dependent chimeric pestivirus
Patent: US 6326137-A 1 04 -DEC-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
           linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12734
176
4
12
3
                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DNA
           DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AR145264 1998 bp D
Sequence 105 from patent US 6211338.
AR179057
Sequence 1 from patent US 6326137
AR179057
                                                                                                                                                                                                                                                   US-09-965-594-22 (1-197) x AR179057 (1-12734)
                                                                                                                                      /organism-"unknown"
2604 c 3295 g
                                     AR179057.1 GI:20220612
                                                                                                                                                                                   3.77e-64
882.50
92.31%
90.26%
86.86%
                                                                         Unclassified
                                                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                  4032 a
                                                        Unknown.
                                                                                                                                                                                                       Percent Similarity:
                                                                  Unknown
                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       953
                                                                                                                                                                                                                                                                     'n
                                                                                                                                                                                                                                                                                                                                                             533
                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                  82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                            42
                                                                                                                                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
         LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISH
                                                                                                                                 source
                                                                                                                                                 BASE COUNT
                                                                                                                                                                                       Pred. No.:
                                                                                                    TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 2
AR145264
LOCUS
                                                                                            AUTHORS
                                                                                   REFERENCE
                                                                                                                       FEATURES
 AR179057
                                                                                                                                                           ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                줨
                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ó
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161
                                                                                                                                                                                                                                                                                                                                          101
                                                                                                                                                                                                                                                                                                                                                                                                             244 TGCGTCAACGGCGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                               41
                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           484 CICICCCCAGGCCIGICICCTACTIGAAGGCTCTICGGGIGGICCACTGCTCTGCCCT
                                                                                                                                                                                                                                                                                                                                                                                               22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                         NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PAT
                                                           1 (bases 1 to 1998)
Malcolm,B.A., faremi.S.Shane., Weber.P.C. and Yao,N.
Malcolm,B.A., faremi.S.Shane., Weber.P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
Patent: US 6211338-A 105 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                      1998
167
15
11
                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AR145268 1998 bp C Sequence 109 from patent US 6211338. AR145268.1 GI:15107135
                                                                                                                                                                       423
                                                                                                                                                                                                                                                                                                                    US-09-965-594-22 (1-197) x AR145264 (1-1998)
                                                                                                                                                        /organism="unknown"
595 c 569 g
GI:15107131
                                                                                                                                                                                                                        1.73e-64
875.50
92.86%
85.20%
                                                                                                                                          .1998
```

```
AR145254
Sequence 95 1
AR145254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown.
Unclassified.
         Unclassified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145254.1
                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown.
                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                     42
                                                                                                                                                                                                                                                                                                                                      83
                                                                                                                                                                                                                                                                                                                                                                                          364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AR145254
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                       Query Match:
                                                                     source
                                                                                                                     ..
Q
ORGANISM
                                                                                    BASE COUNT
                          AUTHORS
                                                  JOURNAL
                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 5
                                                             FEATURES
                                 TITLE
                                                                                              ORIGIN
                                                                                                                       Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                음
                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                 S G
                                                                                                                                                                                                                                                                                                                                  \ddot{o}
                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                 ò
                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     08-AUG-2001
                                                                                                                                                                                                                                                                                                     243
                                                                                                                                                                                                                                                                                                                              244 TGCGTCAACGGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303
                                                                                                                                                                                                                                                                                                                                                       122 ValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                 4.1
                                                                                                                                                                                                                                                                                                                     62 SerileAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                                                                                                                      42 AsnGlnvalGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                   NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAT
                               1 (bases 1 to 1998)
Malcolim, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
protease and NS4A cofactor peptide
Patent: US 6211318-A 109 03-APR-2001;
Location/Qualifiers
1.1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                     1998
166
16
11
3
                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 104 from patent US 6211338. AR145263
                                                                                                                                                                                                (1-1998)
                                                                                           /organism-"unknown"
595 c 569 g
                                                                                                                                                                                                 US-09-965-594-22 (1-197) x AR145268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AR145263.1 GI:15107130
                                                                                                                                     3.09e-64
872.50
92.86%
84.69%
85.88%
                         Jnclassified
                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                       Ø
        Unknown,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                  Unknown
                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  604
                                                                                                                                                                                                                                                    22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                     BASE COUNT
ORIGIN
                ORGANISM
                                                                                                                                        ..
Q
                                                                    JOURNAL
                                          AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 4
AR145263
LOCUS
                                  REFERENCE
KEYWORDS
                                                                            FEATURES
                                                   TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                   ç
                                                                                                                                                                                                                                                                   Ω
                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                    ô
                                                                                                                                                                                                                                                                                                                                   ద
                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                      8
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                           õ
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      543
                                                                                                                                                                                                                                               183
                                                                                                                                                                                                                                                                                                                                                                 243
                                                                                                                                                                                                                                                                                                                                                                                                                                     101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                              41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla
                                                                                                                                                                                                                                                                                              22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                            AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                         NS3
1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and MS4A cofactor peptide
Patent: US 6211338-A 104 03-APR-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                          1998
167
14
12
3
                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N
N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 651 bp
from patent US 6211338.
                                                                                            424
                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                         US-09-965-594-22 (1-197) x AR145263 (1-1998)
                                                                    1. 1998
/organism-"unknown"
1. 596 c 568 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GI:15107121
                                                                                                                                         3.75e~64
871.50
92.35%
85.20%
85.78%
```

of

```
Scores:
                                                                                                                                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                                                                                      122
                                                                                                                                                                                                                                                                                                                                                                                                      484
                                                                                                                                                                                                                                                                                                                                                                                                                                                       182
                                                                                                                                                                                                                                                                                       82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
ACCESSION
                                                 BASE COUNT
ORIGIN
                                                                         Alignment S
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 7
AR145262
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KEYWORDS
SOURCE
                  JOURNAL
                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
  TITLE
                                                                                           Score:
                                                                                                                                                          õ
                                                                                                                                                                         a
                                                                                                                                                                                         ò
                                                                                                                                                                                                        엄
                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                     qq
                                                                                                                                                                                                                                                                                                                       õ
                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                    5
                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                    303
                                                                                                                                                                                                                                                                                                                                                                                                                      543
                                                                                                                                                                                                                                                                                                                                                                                                                                                      603
                                                                                                                                                                                        123
                                                                                                                                                                                                                                                                                                                                                                                                      LeuSerProArgProlleSerTyrLcuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                     162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                      SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                          -----ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                               AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
      Malcolm, B. A., Taremi, S. Shane., Weber, P. C. and Yao, N. Single-chain recombinant complexes of hepatitis C virus protease and NS4A cofactor peptide
Patent: US 6211338-A 95 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 (bases 1 to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlyServalVal11eValGlyArgIleAsnLeuSerGlyAsp---
                                                                                                 651
166
15
11
11
                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 108 from patent US 6211338.
AR145267
AR145267.1 GI:15107134
                                                                  144
                                                                                                                                         Gaps:
                                                                                                                                                         US-09-965-594-22 (1-197) x AR145254 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1998
                                         Location/Qualifiers
1. .651
                                                         /organism="unknown"
187 c 200 g
                                                                                                1.84e-64
868.50
92.82%
85.13%
(bases 1 to 651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
Unclassified.
                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                 120 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AR145267
                                                                                          Alignment Scores:
                                                                                                                                                                                                                                       42
                                                                                                                                                                                                                                                                      62
                                                                                                                                                                                                                                                                                       244
                                                                                                                                                                                                                                                                                                                                                                                                      142
                                                                                                                                                                                                                                                                                                                                                                                                                    4 B 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                 Query Match:
                                                  source
                                                                 BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
VERSION
                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 6
AR145267
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ą
                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                   5
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                        ò
                                                                                                                                                                                       g
                                                                                                                                                                                                        à
                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                      ð
                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                     õ
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                               41
                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                                           7
                                                                                                                                                                                                                                                                           ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTGCTGGTGGTCCACTGCTCTGCCCT
                                                                                                                                                                                                                                                          GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                 42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (bases 1 to 1998)
Malcolm, B.A., Taremi, S. Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus
protease and NS4A cofactor peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                  1998
166
15
12
3
                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
Single-chain recombinant complexes of protease and NS4A cofactor peptide Patent: US 6211338-A 108 03-APR-2001; Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 103 from patent US 6211338. AR145262
                                                                                 424
                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                 (1-1998)
                                                                   /organism="unknown"
596 c 568 g
                                                                                   ь
                                                                                                                                                                                                                                 US-09-965-594-22 (1-197) x AR145267
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AR145262.1 GI:15107129
                                                                                                                                  6.72e-64
868.50
92.35%
84.69%
85.48%
                                                      1. .1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unclassified
                                                                                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                 ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown
```

source 1. COUNT 120 a IN		Qy 5 GlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAlaTyr 21	GGTAGTAT	22 AlaGlnThrArgGlyGluGlnGlnGlnLysThrSerHisThrGlyArgAspLys :::	124 TCCCAACAGGGGGCCTACTTGGTTGCAAGAAGACTAGCCTTACAGGCCGGGACAAG	Oy 42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr 61	10. ANCAGGICANGGGANDANGIICAGGGANGCACACACACACATICTIGGGGANC 24 62 SerileAsnGlyValleuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81	Db 244 TGCGTCAACGGCGTGTTGGACCGTTTACCATGGTGGTGGCTCAAAGACCTTAGCCGGC 303	Oy 82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101	304 CCAAAGGGGCCAATCACCCAGATGTACACTAATGTGGACCAGGACCTCGTCGGCTGGCAG	Qy 102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121	Qy 122 ValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySerLeu 141	424 GTCACGAGACATGCTGACGTCATTCCGGTGCGCCGGCGGCGACAGTAGGGGGAGCCTG	UY 142 DEMONSTRATE DEMONSTRATE DE LES SELENTATION DE LA DEMONSTRATE DE LA DESENTATION DE LA DEMONSTRATE DEL DEMONSTRATE DEL DEMONSTRATE DE LA DELLA DEL DEMONSTRATE DEL DEMONSTRATE DE LA DEMONSTRATE DE LA DEMONSTRATE DEL DEMONSTRATE DE LA DEMONSTRATE DEL DEMONSTRATE DEL DESCRIPTOR DE LA DEMONSTRATE DEL DEMONSTRATE DEL DEMONSTRATE DEL DESCRIPTOR DEL DESCRIPTOR DEL DESCRIPTOR DEL DESCRIPTOR DEL DESCRIPTOR DE LA DESCRIPTOR DE LA DESCRIPTOR DEL DEL DEL DESCRIPTOR DEL DEL DESCRIPTOR DEL DESCRIPTOR DEL DEL DESCRIPTOR DEL	AlaGiyHisAlaValGiyIlePheArgAlaAlaValSerThrArgGiyValAlaLysAla	544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCGGG	Oy 182 ValAspPhelleProValGluSerLeuGluThrThrMetArgSer 196 		AR142423 AR142423 AR142423 AR142523 ACCUSS DEFINITION Sequence 94 from patent US 6211338. ACCESSION AR145253 VERSION AR145253.1 GI:15107120 KEYWORDS SOURCE Unknown.	Unclassified. 1 (bases 1 to 651) Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.	TILE SIGGLE-Chain recombinant complexes of hepatilis C virus NS3 protease and NS4A cofactor peptide JOURNAL Patent: US 6211338-A 94 03-APR-2001; FRATIDES	source 1651 /organism="unknown"
JOURNAL Patent: US 6211338-A 103 03-APR-2001; FEATURES Location/Qualifiers 1. 1998 Source /organism="unknown" BASE COUNT 410 a 596 c 568 g 424 t ORIGIN	Alignment Scores: 8.16e-64	US-09-965-594-22 (1-197) x AR145262 (1-1998)	Oy 5 GlySerValValIleValGlyArglleAsnLeuSerGlyAspThrAlaTyr 21	64 GGITCTGIIGHIHHHHHHHHHHHHHHHHHHHHHHHHHHHH	22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys	Db 124 TCCCAACAGACGGGGCCTACTTGGTTGCAAGATCACTTACAGGCCGGGACAAG 183	184 AACCAGGTCGAGGGAGGTTCAGGTGGCTTTCCACCGCAACACAATCTTCCTGGCGACC	Qy 62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81	Db 244 TGCGTCAACGGCGTGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303	82 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln	Db 304 CCAAAGGGCCAATCACCCGAGTGTACACTAATGTGGACCAGGACCTGGTGGCTGGC	Db 364 GCGCCCCCGGGGCGCGTTCCTTGACACCATGCACCTGGGCAGCTCAGACCTTTACTTG 423	Oy 122 ValThrArgHisAlaAspVallleProValArgArgGlyAspSerArgGlySerLeu 141	142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro		162 AlaciyhisAlaValGiyilePheArgAlaAlaValSerThrArgGiyValAlaLysAla	3TTGCC 197		RESULT B AR145258 LOCUS DEFINITION Sequence 99 from patent US 6211338. ACCESSION AR145258 REYMORDS REYMORDS	_	ACTURED 1 (Dates 1 LO 01) AUTHORS Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N. TITLE Single-chain recombinant complexes of hepatitis C virus NS3 Drotease and NS4A cofactor peptide	JOURNAL Patent: US 6211338-A 99 03-APR-2001; FEATURES Location/Qualifiers

us-09-965-594-22.rge

```
Alignment Scores:
       Alignment Scores:
Pred. No.:
                                                                                'n
                                                                                                                                                                                                                                                                                                                                                  182
                                                                                                                                         42
                                                                                                                                                                      62
                                                                                                                                                                                                  83
                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
DEFINITION
ACCESSION
VERSION
                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                              RESULT 11
AR145261
                                                                                                                                                                                                                                                                                                                                                                                                                         KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                        Score:
                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                à
                                                                                             qq
                                                                                                            ö
                                                                                                                          qq
                                                                                                                                         δ
                                                                                                                                                     qq
                                                                                                                                                                     ò
                                                                                                                                                                                   g
                                                                                                                                                                                                  ò
                                                                                                                                                                                                              q
                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                    셤
                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                          08-AUG-2001
                                                                                                                                                                                                               423
                                                                                                                                                                                                                                                                        243
                                                                                                                                                                                                                                             102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                       61
                                                                                                                          41
                                                                                              21
                                                                                                                                                     22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown.
Unclassified.
1 (bases 1 to 1998)
Malcolm,B.A., Taremi,S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
Single-chain recombinant complexes of hepatitis C virus NS3
Protease and NSA cofactor peptide
Patent: US 6211338-A 107 03-APR-2001;
Location/Qualifiers
1. 1998
                                                                                                                                                                                                                                                                                                                                                               182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                              648
                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                                                                                                                                                                                                                                       5 GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp--
                             651
166
14
12
3
                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 107 from patent US 6211338. ARI45266
ARI45266.1 GI:15107133
145
                                                                Gaps:
                                                                               US-09-965-594-22 (1-197) x AR145253 (1-651)
Q
199
                           4e-64
864.50
92.31%
85.13%
85.09%
U
                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown.
                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                             ..
₽
                                                                                                                                                                                                                                                                                                                                                                                           RESULT 10
AR145266
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
      DRIGIN
                             Pred.
                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                              õ
                                                                                                         g
                                                                                                                          ò
                                                                                                                                       g
                                                                                                                                                     ò
                                                                                                                                                                   d
                                                                                                                                                                                   à
                                                                                                                                                                                               엄
                                                                                                                                                                                                               õ
                                                                                                                                                                                                                             8
                                                                                                                                                                                                                                         οχ
                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                  ò
```

```
08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                       543
                                                                                                                                                                                                                                                                                                          121
                                                                                                                                                                                                                                                                                                                                                                  423
                                                                                                                                                                                                                                                                                                                                                                                                                                       161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                243
                                                                                                                                                                                                                                                                        303
                                                                                                                                                                                                                                                                                              101
                                                                                                                    4.1
                                                                                                                                                                                                    61
                                                                                                                                                                                                                                                 81
                                                                                                                                                                                                                                                                                                                                                                                                                                   244 TGCGTCAACGGCGTGTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArqGlyValAlaLysAla
                                                                                                                                                                                                               ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                        GlySerValValIleValGlyArglleAsnLeuSerGlyAsp.-----ThrAlaTyr
                                                                                                                                                                                                   AsnglnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (bases I to 1998)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Patolease and NASA coffector peptide
Patent: US 6211338-A 102 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ValAspPhelleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             linear
1998
165
16
12
3
            Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sequence 102 from patent US 6211338.
ARI45261
ARI45261.1 GI:15107128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              425
                                              Indels:
                                                                                 US-09-965-594-22 (1-197) x AR145266 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
1. .1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /organism="unknown"
597 c 567 g
1.46e-63
864.50
92.35%
84.18%
85.09%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unclassified
                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              409 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown
```

Percent Similarity: 91.84% Conservative: 14 Best Local Similarity: 84.69% Mismatches: 13 Query Match: 84.99% Indels: 3 DB: 6 Gaps: 1 US-09-965-594-22 (1-197) x AR145269 (1-2016)	Oy 5 GlySerValValIleValGlyArgileAsnLeuSerGlyAspThrAlaTyr 21	20 mc.nost.consecons.conse	•	AREALT 13 AREAGLA B. LOCUS LOCUS LOCUS AP268278 LOCUS AP268278 ACCESSION VERSION AP268278 AP2	y
Pred. No.: 1.77e-63 Length: 1998 Score: 863.50 Matches: 166 Percent Similarity: 91.84% Conservative: 14 Best Local Similarity: 84.69% Mismatches: 13 Query Match: 6 Gaps: 1	US-09-965-594-22 (1-197) x AR145261 (1-1998) OY 5 GlySerValValIleValGlyArglleAsnLeuSerGlyAspThralaTyr 21	184 AACCAGGTCAAGGAGAGTTCAGGTGTTTCCACCGCAACACATTCCTCCTGGCGACC 62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIlealaSer 111111111111111111111111111111111111	Db 364 GGGCCCCGGGGGCGGTTCCTTGACACCATGCACCTGTGCACCTTTACTTG 423 Qy 122 ValThrArgHisAlaAspvallieProValargArgArgGlyAspSerArgGlySerLeu 141	SULT 12 CUS	source 12016 Assection And And And And And And And And And An

```
EGEPEKTACTENEES WEEL VEGINGE TO THE STATE OF THE SERVED EVERTACTER TO THE SERVED TO THE S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GKTEKENTODELTHRINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGENEERINGTONG TEGTORAMERGENEERINGTONG TEGTORAMERGENEERINGTONG TEGTORAMERGENEERINGTONG TEGTORAMERGENEERINGTONG TEGTORAMERGENEERINGTONG THE ROBERT TANGENEERINGTONG THE STOREN THE STO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   QDLVGWPAPQGSRSLIPCTCGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRPISYLKG
SSGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
SYSDIKEEGATKKKIQKPDRLERGKMKIVPKESEKDSKTKPPDAIIVVEGVKYQVRKK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMKAFKNVGGKLTKVEESGPPLCRNRPGRGPVNYRVTKYYDDNLREIKPVAKLEGOVE
FYKOVTAK IDYSKGKMLLAPDKWEVENBYTYELAKTYGOVERKAYLGOBPNHRALV
ERDCATITKNTVQFLKMKKGGETYDLTISNIJRLIELVHRNNLEEKEIPTAYTYML
AYTEVNEDVGTIKPVLGERVIPDPVVDINLQPEVQVDTSEVGITIIGRETLATTGYTP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VLEKVEPDASDNQNSVKIGLDEGNYPGPGIQTHTLTEEIHNRDARPFIMILGSRNSIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NRAKTARNI NLYTGNDPREI RDLMAAGRML VVALRDVDPELSEWVDFKGTFLDREALE
ALSLGQPKPKQVTKEAVRNL I EQKKDVEI PNWFASDDPVFLEVALKNDKYY LVGDVGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VKDQAKALGATDOTRI IKEVGSRTYAMKLSSWFLQASNKQMSLTPLFEELLLRCPPAT
KSNKGHMASAYQLAQGNWEPLGCGVHLGTIPARRVKIHPYEAYLKLKDFIEEEEKKPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VKDTVIREHNKWILKKIRFQGNLNTKKMLNPGKLSEQLDREGRKRNIYNHQIGTIMSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /GEVTWEQLEAGINRKGAAGFLEKKNIGEVLDSEKHLVEQLVRDLKAGRKIKYYETAI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PKNEKRDVSDDWQAGDLVVEKRPRVIQYPEAKTRLAITKVMYNWVKQQPVVIPGYEGK
TPLFNIFDKVRKEMDSFNEPVAVSFDTKAMDTQVTSKDLQLIGEIQKYYYKKEMHKFI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AGIRLEKLPIVRAQTDTKTFHEAIRDKIDKSENRQNPELHNKLLEIFHTIAQPTLKHT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /translation="MELNTNEGSGSVVIVGRIVLSGSGSITACAQQTRGLLGCKITSL
IGRDKNQVEGEVQIVSTATQTFLATCINGVCWIVYHGAGTRTIASPKGPVIQMYINVD
/organism~"Pestivirus type 1"
                                                                                                                                                                                                                                                                                                                                                                                                                                           /product="polyprotein"
/protein_id="AAF82566.1"
/db_xref-"G1:9049957"
                                                                   /mol_type="genomic RNA"
/db_xref="taxon:11099"
                                                                                                                                                                                                                                                                                                                                                                               /codon_start=1
                                                                                                                                                                                                                                                                                                           .12508
                                                                                                                                                                                                                                        .385
```

```
IEFCSHIPVPVRWSDMTSSHAAGNINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILLINGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGSTATILNGS
DTITDHMTEVPVITADGEVYIRNGQRGSGQPDTSAGNSMLNVLTMMYAFCESTGVPYK
SFNRVARIHVCGDDGFLITEKGLGLKFANKGMQILHEAGKPQKITEGEKMKVAYRFED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               712
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     832
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         892
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             952
                                                                                                                                                                                                                                                                                                                                                                                                                                                               532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                122 ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                      GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       833 CTTTCGCCCCGGCCCATTTCCTACTAAAGGCTCCTCTGGGGGTCCGCTGTTGTGCCCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (bases 1 to 651)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chain recombinant complexes of hepatitis C virus NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA
                                                                                                                                                     g 2802 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     51 bp
6211338.
                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-22 (1-197) x AF268278 (1-12734)
                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 551
Sequence 98 from patent US 62
AR145257
AR145257.1 GI:15107124
                                                                                                                                                   3293
                                                                                                                                                                                                                                       1.83e-62
862.50
91.28%
88.72%
84.89%
                                                                                                                                                        U
                                                                                                                                                   2608
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unclassified
                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                        ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unknown.
                                                                                                                                                     4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Unknown
                                                                                                                                                                                                                     Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                      'n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             953
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
ACCESSION
                                                                                                                               3'UTR
BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                     Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 14
AR145257
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      엄
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

g

```
ಗ
                                               Alignment Scores:
                                                                                                                                                                                                                                               82
                                                                                                                                                                                                                                                                                                                                                                                                    182
                                                                                                                                                                                                                                                                                                                                                                                                                  604
                                                                                    Query Match:
DB:
                                                       ..
₽
                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                          Search cor
Job time
 FEATURES
                               ORIGIN
                                                              Score:
                                                       Pred.
                                                                                                                                        a
                                                                                                                                                                                                                                                         Q
                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                         ö
                                                                                                                                                      Ω
                                                                                                                                                                    g
                                                                                                                                                                                    ò
                                                                                                                                                                                               g
                                                                                                                                                                                                                õ
                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                08-AUG-2001
                                                                                                                                                                                                                                                                                                                       363
                                                                                                                                                                                                                                                                                                  543
                                                                                                                                                                                                                                                                                                                                                                                            243
                                                                                                                                                                                                                                               303
                                                                                                                                                                                                                                                             101
                                                                                                                                                                                                                                                                                           121
                                                                                                                                                                                                                                                                                                                                                      142 LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                     162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                  61
                                                                                                                                                                                                                                81
                                                                                                                                                                    41
                                                                                                                                                                                                                                                            AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                  42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                         AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 (bases 1 to 651)
Malcolm.B.A., Taremi.S.Shane., Weber,P.C. and Yao,N.
Single-chain recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 93 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAT
                                                                                                                                                                                                                                                                                                                                                                                                                            182 ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                    651
165
15
12
                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                      u
protease and NS4A cofactor peptide
Patent: US 6211338-A 98 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 93 from patent US 6211338.
AR145252
                                      145
                                                                                                         Gaps:
                                                                                                                        US-09-965-594-22 (1-197) x AR145257 (1-651)
              Location/Qualifiers
1. .651
                            /organism="unknown"
188 c 199 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GI:15107119
                                                                   7.16e-64
861.50
92.31%
84.62%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unclassified
                                                                              Percent Similarity:
Best Local Similarity: |
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR145252.1
                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                                                             Alignment Scores:
                                                                                                                                                                    22
                                                                                                                                                                                                                184
                                                                                                                                                                                                                                                             82
                                                                                                                                                                                                                                                                            304
                                                                                                                                                                                                                                                                                           102
                                                                                                                                                                                                                                                                                                                        122
                                                                                                                                                                                                                                                                                                                                        124
                                                                                                                                                                                                                                                                                                                                                                      484
                                                                                                                                                                                                                                                                                                                                                                                                                                 604
                                                                                                                                                                                                                                                                                                                                                                                                                                                        AR145252
LOCUS
DEFINITION
                        source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
                                                                    Pred. No.:
                                      BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
       JOURNAL
                FEATURES
                                              DRIGIN
                                                                             Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                            q
                                                                                                                                                                                                                               οÿ
                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                        5
                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                        ò
                                                                                                                                                 셤
                                                                                                                                                                    ç
                                                                                                                                                                                셤
                                                                                                                                                                                             ò
```

```
483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  543
                                                                                                                                                                                                     183
                                                                                                                                                                                                                                                                                                           243
                                                                                                                                                                                                                                                                                                                                                        303
                                                                                                                                                                                                                                                                                                                                                                                101
                                                                                                                                                                                                                                                                                                                                                                                            121
                                                                                                                                                                                                                                                                                                                                                                                                                                                     423
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122 ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                603
                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                       41
                                                                                                                                                                                                                                                                                                                                  83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                          22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLySThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                    ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                                                                                                                                                         GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                   42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GTGGACTTTGTGCCCGTAGAGTCCATGGAAACTACTATGCGGTCT
                                                                                 651
165
15
12
3
                                                                                                       Conservative:
Mismatches:
                                                                                 Length:
Matches:
                                     145
                                                                                                                                Indels:
                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            completed: August 31, 2003, 00:46:41
ne: 2569.57 secs
                                                                                                                                                                   US-09-965-594-22 (1-197) x AR145252 (1-651)
Location/Qualifiers
1..651
                        /organism="unknown"
188 c 199 q
                                     ρ
                                                                               8.69e-64
860.50
92.31%
84.62%
84.69%
                                                                                                         Percent Similarity:
Best Local Similarity:
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

```
// Search time 182.939 Seconds
(without alignments)
2906.924 Million cell updates/sec
                                                                                                                                                    1 MKKKGSVVIVGRINLSGDTA......VAKAVDFIPVESLETIMRSP 197
- nucleic search, using frame_plus_p2n model
                                     August 30, 2003, 19:13:57
                                                                                                         US-09-965-594-22
1016
                                                                                                                                                                                              BLOSUM62
                                                                                                                                  Perfect score:
                                                                                                                                                                                              Scoring table:
OM protein
                                                                                                                                                       Sequence:
                                       Run on:
```

0.5 Xgapop 10.0 , Xgapext Ygapop 10.0 , Ygapext Fgapop 6.0 , Fgapext Delop 6.0 , Delext

5105512 of hits satisfying chosen parameters: Total number

2552756 segs, 1349719017 residues

Searched:

length: 0 length: 2000000000 Minimum DB seq Maximum DB seq

summaries Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 su

Command line parameters:
-WODEL-frame+_p2n.model -DEV-x1p
-Q-Copn2_1/USPTO_spool/US09955594/runat_29082003_151918_28302/app_query.fasta_1.2872
-Q-Copn2_1/USPTO_spool/US09955594/runat_29082003_151918_28302/app_query.fasta_1.2872
-Q-Copn2_1/USPTO_spool/US099555494/runat_29082003_151918_28302/app_query.fasta_1.2872
-DB-N_Geneseq_19Jund3 -OPMT-fasta_P -USFTIX-spool/USBMC2 -TRANS-human40.cd1
-LIST-A-COPETTO_ -UNITS-bits -STRF-1 -END-1 -MARRIX-bloo -THR_NIN-O -ALIGN-15
-MODE-LOCAL -OUTEN-PDO -NORM-ext -HEAPSIZE-500 -MINLEN-O -MAXEN-200000000
-USER-USG995554_@CCN_1 _1142_@tunat_29082003 151918_28302 -NPV-6 -ICPU-3
-NO_MMAP -LANGEQUERY -NEG_SCORES- O -WAIT -DSPBLOCK=100 -LONGLOG
-EVEV_TIMEOOT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELDEXT-7

N_Geneseq_19Jun03:*

Database :

1: /SIDS1/gcgdata/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseqn-embl/NA1981.DAT:*
4: /SIDS1/gcgdata/geneseqn-embl/NA1981.DAT:*
5: /SIDS1/gcgdata/geneseqn-embl/NA1982.DAT:*
6: /SIDS1/gcgdata/geneseqn-embl/NA1986.DAT:*
7: /SIDS1/gcgdata/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseqn-embl/NA1980.DAT:*
8: /SIDS1/gcgdata/geneseqn-embl/NA1980.DAT:*
9: /SIDS1/gcgdata/geneseqn-embl/NA1980.DAT:*
10: /SIDS1/gcgdata/geneseqn-embl/NA1991.DAT:*
11: /SIDS1/gcgdata/geneseqn-embl/NA1991.DAT:*
12: /SIDS1/gcgdata/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseqn-embl/NA1991.DAT:*
14: /SIDS1/gcgdata/geneseqn-embl/NA1991.DAT:*
15: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1991.DAT:*
16: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1991.DAT:*
17: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1991.DAT:*
18: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1991.DAT:*
19: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
19: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
20: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
22: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
23: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
24: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
25: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
26: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
27: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
28: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
28: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:*
29: /SIDS1/gcgdata/geneseqn/geneseqn-embl/NA1999.DAT:* : !

/SIDS1/gcgdata/geneseg/genesegn-emb1/NA2001A.DAT:*/SIDS1/gcgdata/geneseg-emb1/NA2001B.DAT:*/SIDS1/gcgdata/geneseg-emb1/NA2001B.DAT:*/SIDS1/gcgdata/geneseg-emb1/NA2002.DAT:*/SIDS1/gcgdata/geneseg-geneseg-emb1/NA2003.DAT:*

Pred. No. 1s the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

### ### ### ### ### ### ### ### ### ##	+[0		: : :				
1 1016 100.0 994 21 AAA73334 Hepatitis C VIT	No.	COL	Match		DB		Description
2 995 97. 594 21 AAA73333 Hepatitis C VIT		10	100.	594	21	333	Hepatitis C viru
4 986 97.9 594 21 AAA73332 Hepatitis C VIT AAA73334 Hepatitis C VIT 6 96 96 94.8 594 21 AAA7333 Hepatitis C VIT 6 96 96 94.8 594 21 AAA7332 Hepatitis C VIT 9 92 91.8 594 21 AAA7332 Hepatitis C VIT 9 92 91.8 588 21 AAA7332 Hepatitis C VIT 9 92 91.8 588 21 AAA7332 Hepatitis C VIT 9 92 91.8 588 21 AAA7332 Hepatitis C VIT 9 92 91.8 588 21 AAA7332 Hepatitis C VIT 9 92 91.8 58.9 1998 20 AAX8035 Hepatitis C VIT 9 92 91.8 58.9 1998 20 AAX8035 Hey MCV N84-N83 CO VIT 9 92 91.8 58.9 1998 20 AAX8035 HCV N84-N83 CO VIT 9 92 92 92 92 92 92 92 92 92 92 92 92 9	7	0	<u>.</u>	594	77	333	itis C
4 986 96.5 84 21 AAA73331 Hepatitis C vir 7 929 92.1 594 21 AAA7333 Hepatitis C vir 7 929 92.1 584 21 AAA7332 Hepatitis C vir 7 929 92.1 584 21 AAA7323 Hepatitis C vir 929 92.1 588 21 AAA7322 Hepatitis C vir 929 92.2 8.8 8 92.1 588 21 AAA7322 Hepatitis C vir 929 92.2 8.8 9123 92.2 AAX8035 Hepatitis C vir 824.83 coll 872.5 86.2 1939 20 AAX8035 Hey 824.83 coll 872.5 86.2 1939 20 AAX8035 Hey 824.83 coll 872.5 85.4 1938 20 AAX8035 Hey 824.83 coll 885.5 85.4 651 20 AAX8035 Hey 824.83 coll 885.5 85.4 651 20 AAX8035 Hey 824.83 coll 885.5 85.4 651 20 AAX8035 Hey 824.83 coll 886.5 85.1 651 20 AAX8035 Hey 824.83 coll 886.5 84.7 1998 20 AAX8035 Hey 824.83 coll 886.5 84.1 651 20 AAX8035 Hey 826.8 84.1 651 20 AAX8035 Hey 824.5 Hey 824.5 84.1 651 20 AAX8035 Hey 824.5 84.	m ·	6	ζ.	594	21	AAA73332	itis C
Secondary Seco	→ (∞、	٠.	594	77	AAA73331	O
10 10 10 10 10 10 10 10	n v	90	ė	450	77	AAA73330 ***73336	o c
8 90.5 86.8 21 AAAV3325 Hebritis C vir 8 90.5 86.9 12734 AAAV3325 Hebritis C vir 9 875.5 86.9 12734 AAAV3325 Hebritis C vir 1 877.5 86.9 12734 AAAV3325 Hebritis C vir 1 877.5 86.9 12734 AAAV3325 Hev N844-N83 co 1 877.5 86.9 12734 AAAV380354 Hev N844-N83 co 1 865.5 85.5 65.1 20 AAX80355 Hev N844-N83 co 1 865.5 85.2 1998 20 AAX80355 Hev N844-N83 co 1 865.5 85.1 1998 20 AAX80355 Hev N844-N83 co 1 865.5 85.1 1998 20 AAX80354 Hev N844-N83 co 1 865.5 85.1 150.2 AAX80354 Hev N844-N83 co 1 865.5 85.1 1998 20 AAX80355 Hev N844-N83 co 1 865.5 85.1 1998 20 AAX80355 Hev N844-N83 co 1 865.5 85.1 1998 20 AAX80355 Hev N844-N83 co 1 865.5 85.1 1998 20 AAX80355 Hev N844-N83 co 1 865.5 86.1 1998 20 AAX80355 Hev N844-N83 co 2 865.5 84.4 65.0 20 AAX80355 Hev N844-N83 co 3 855.5 84.4 65.0 20 AAX80355 Hev N844-N83 co 4 8 8 8 8 10 65.0 20 AAX80345 Hev N844-N83 co 5 8 8 8 8 8 20 AAX80345 Hev N844-N83 co 8 8 8 8 8 8 8 8 8	2	へへ	, .	288	217	AAA73329	ى ر
9 8825 86.9 12734 24 AAR895615 Chimnerte BryDVAR 11 875.5 86.9 12734 24 AAR89555 Chimnerte BryDVAR 11 875.5 86.2 1998 20 AAR89355 Chimnerte BryDVAR 12 86.5 85.9 1998 20 AAR89355 Chimnerte BryDVAR 12 86.5 85.9 1998 20 AAR89356 Chimnerte BryDVAR 13 coll 16 86.5 85.1 651 20 AAR89356 Chimnerte BryDVAR 14 86.5 85.1 651 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 64.0 650 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 64.0 650 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 64.0 650 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 64.0 650 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 64.0 650 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 64.0 650 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 64.0 650 20 AAR89356 Chimnerte BryDVAR 18 864.5 85.1 690 19 AAR89356 Chimnerte BryDVAR 18 864.5 83.1 7310 10 AAR99336 Combined Open r Companied Chimnerte BryDVAR 18 864.5 83.1 7310 10 AAR99336 Combined Open r Combined Open r Combined Open r Companied Chimnerte BryDVAR 18 864.5 83.1 7330 10 AAR99336 Combined Open r Combined Ope	- 00	9	· 🕉	588	21	AAA73328	יטי
10 875.5 86.7 1998 20 AAX80355 HCV NSAA-NS3 CO 12 871.5 85.8 1998 20 AAX80355 HCV NSAA-NS3 CO 14 868.5 85.9 1998 20 AAX80355 HCV NSAA-NS3 CO 15 867.5 85.4 1998 20 AAX80355 HCV NSAA-NS3 CO 16 865.5 85.4 1998 20 AAX80355 HCV NSAA-NS3 CO 17 864.5 85.1 1998 20 AAX80355 HCV NSAA-NS3 CO 18 865.5 85.1 1998 20 AAX80355 HCV NSAA-NS3 CO 22 861.5 85.1 1998 20 AAX80355 HCV NSAA-NS3 CO 23 861.5 84.7 1998 20 AAX80355 HCV NSAA-NS3 CO 24 860.5 84.7 1651 20 AAX80356 HCV NSAA-NS3 CO 25 860.5 84.7 1651 20 AAX80356 HCV NSAA-NS3 CO 26 860.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 27 860.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 28 857.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 29 857.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 20 867.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 20 867.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 21 867.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 28 857.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 29 857.5 84.4 651 20 AAX80356 HCV NSAA-NS3 CO 30 865.5 84.1 651 20 AAX80356 HCV NSAA-NS3 CO 31 863.8 84.1 651 20 AAX80356 HCV NSAA-NS3 CO 32 857.5 84.1 651 20 AAX80356 HCV NSAA-NS3 CO 34 881.8 83.1 662.2 1AAX82258 HCV NSAA-NS3 CO 36 848.5 83.5 12880 1AAX82358 HCV NSAA-NS3 CO 37 848.5 83.5 12880 1AAX82358 HCV NSAA-NS3 CO 44 844.5 83.1 7310 10 AAN92106 COmbined Open r 44 844.5 83.1 7310 10 AAN92106 COmbined Open r 45 844.5 83.1 7310 10 AAN92106 COMBINED OPEN R 44 844.5 83.1 7310 10 AAN92106 COMBINED OPEN R 45 844.5 83.1 7310 10 AAN92106 COMBINED OPEN R 46 844.5 83.1 7310 10 AAN92106 COMBINED OPEN R 46 844.5 83.1 7310 10 AAN92106 COMBINED OPEN R 47 844.5 83.1 7310 10 AAN92106 COMBINED OPEN R 48 844.5 83.1 7310 10 AAN92106 COMBINED OPEN	σ,	٠.	٠.	273	24	ABA95615	BVDV
12 07.1 07.2 07.2 0.0 13 668.5 85.9 1398 20 0.0 14 668.5 85.1 1598 20 0.0 15 665.5 85.4 1998 20 0.0 15 665.5 85.4 1998 20 0.0 16 685.5 85.4 1998 20 0.0 17 864.5 85.1 1998 20 0.0 18 864.5 85.1 1998 20 0.0 19 864.5 85.1 1998 20 0.0 10 864.5 85.1 1998 20 0.0 10 864.5 85.1 1998 20 0.0 11 864.5 85.1 1998 20 0.0 12 865.5 85.1 1998 20 0.0 13 865.5 85.1 1998 20 0.0 14 865.5 85.1 1998 20 0.0 15 865.5 85.1 1998 20 0.0 16 865.5 85.1 1998 20 0.0 17 864.5 85.1 1998 20 0.0 18 865.5 85.1 1998 20 0.0 18 865.5 85.1 1998 20 0.0 19 864.5 85.1 1998 20 0.0 19 864.5 85.1 1998 20 0.0 10 864.5 85.1 1998 20 0.0 10 864.5 85.1 1998 20 0.0 11 864.5 85.1 1998 20 0.0 12 865.5 86.5 86.5 86.5 86.5 13 865.5 86.5 86.5 86.5 86.5 14 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 15 865.5 86.5 86.5 86.5 86.5 16 865.5 86.5 86.5 86.5 86.5 17 19 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 18 86.5 86.5 86.5 86.5 18	10		٠ س	500	200	AAX80355	NS4A-NS3
13 668 5 651 6 5 651 10 AAX80346 HCV NSTA-NS3 CONSTA-NS3 CONSTANT 14 668 5 651 6 5 651 20 AAX80358 HCV NSTA-NS3 CONSTANT 15 665 5 651 20 AAX80346 HCV NSTA-NS3 CONSTANT 16 664 5 651 20 AAX80346 HCV NSTA-NS3 CONSTANT 17 664 5 651 1 651 20 AAX80346 HCV NSTA-NS3 CONSTANT 18 664 5 651 1 651 20 AAX80346 HCV NSTA-NS3 CONSTANT 18 664 5 651 1 651 20 AAX80356 HCV NSTA-NS3 CONSTANT 18 663 5 650 1 1098 20 AAX80356 HCV NSTA-NS3 CONSTANT 18 663 5 664 5 651 20 AAX80346 HCV NSTA-NS3 CONSTANT 18 663 5 644 5 651 20 AAX80346 HCV NSTA-NS3 CONSTANT 18 660 5 644 7 648 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 660 5 644 7 648 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 660 5 644 6 648 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 660 5 644 6 648 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 664 5 654 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 664 5 648 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 664 5 648 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 644 5 644 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 648 5 644 6 651 20 AAX80363 HCV NSTA-NS3 CONSTANT 18 648 5	12	,	'n.	ממת	0 0	AAX80354	NS4A-NSS NS4A-NSS
14 666.5 85.5 1998 20 AAX80358 HCV N84A-N83 co. 15 667.5 85.2 1998 20 AAX80353 HCV N84A-N83 co. 16 665.5 85.2 651 20 AAX80349 HCV N84A-N83 co. 17 664.5 85.1 612 25 AAX80349 HCV N84A-N83 co. 18 664.5 85.1 612 25 AAX80354 HCV N84A-N83 co. 19 664.5 85.1 1998 20 AAX80355 HCV N84A-N83 co. 20 663.5 85.0 2013 20 AAX80355 HCV N84A-N83 co. 21 860.5 84.7 651 20 AAX80356 HCV N84A-N83 co. 22 860.5 84.7 1998 20 AAX80356 HCV N84A-N83 co. 23 860.5 84.7 1998 20 AAX80356 HCV N84A-N83 co. 24 860.5 84.4 651 20 AAX80365 HCV N84A-N83 co. 25 860.5 84.4 651 20 AAX80365 HCV N84A-N83 co. 26 860.5 84.4 651 20 AAX80365 HCV N84A-N83 co. 27 860.5 84.4 651 20 AAX80365 HCV N84A-N83 co. 28 860.5 84.4 651 20 AAX80365 HCV N84A-N83 co. 29 857.5 84.4 651 20 AAX80365 HCV N84A-N83 co. 29 857.5 84.4 651 20 AAX80365 HCV N84A-N83 co. 29 857.5 84.0 651 20 AAX80365 HCV N84A-N83 co. 29 857.5 84.0 651 20 AAX80365 HCV N84A-N83 co. 29 857.5 84.1 648 20 AAX80365 HCV N84A-N83 co. 20 866.5 84.1 648 20 AAX80365 HCV N84A-N83 co. 20 866.5 84.1 648 20 AAX80365 HCV N84A-N83 co. 20 866.5 84.1 648 20 AAX20259 HCV N84A-N83 co. 20 866.5 84.1 648 20 AAX20259 HCV N84A-N83 co. 20 866.5 84.2 83.5 12980 24 AAX30259 HCV N84A-N83 co. 20 866.5 84.5 83.5 12980 24 AAX30259 HCV N84A-N83 co. 20 866.5 84.5 83.5 12980 24 AAX30259 HCV N84A-N83 co. 20 866.5 84.5 83.1 7310 10 AAN93030 Combined open r. 20 866.5 84.5 83.1 7310 10 AAN93040 Combined open r. 20 866.5 84.5 83.1 7310 10 AAN93040 Combined open r. 20 866.5 84.5 83.1 7310 10 AAN93040 Combined open r. 20 866.5 84.5 84.5 84.5 84.5 84.5 84.5 84.5 84.5 84.5 84.5 84.5	13			55	202	AAX80345	NS4A-NS3
15 865.5 85.2 1998 20 AAX80345 HCW NG4A-NS3 CO	14	œ i	ر ا	66	20	AAX80358	NS4A-NS3
10 664.5 85.1 651.2 20 AAX80349 AAX80340	15	۲.	'n.	9	20	AAX80353	NS4A-NS3
B 64.5 85.1 651 20 AAX80344 HCV NS4A-NS3 CO	97		'nv	450	2 5	AAX80349 ABX14706	NS4A-NS3
19 864.5 85.1 1998 20 AAX80352 HCV NS4A-NS3 CO 865.5 85.0 1998 20 AAX80352 HCV NS4A-NS3 CO 865.5 85.0 1998 20 AAX80352 HCV NS4A-NS3 CO 860.5 84.7 651 20 AAX80364 HCV NS4A-NS3 CO 860.5 84.7 1998 20 AAX80365 HCV NS4A-NS3 CO 860.5 84.4 654 20 AAX80365 HCV NS4A-NS3 CO 865.5 84.4 651 20 AAX80365 HCV NS4A-NS3 CO 865.5 84.4 651 20 AAX80363 HCV NS4A-NS3 CO 865.5 84.4 651 20 AAX80363 HCV NS4A-NS3 CO 865.5 84.4 651 20 AAX80364 HCV NS4A-NS3 CO 865.5 84.6 650 20 AAX80362 HCV NS4A-NS3 CO 865.5 84.0 650 20 AAX80364 HCV NS4A-NS3 CO 865.5 84.0 650 20 AAX80362 HCV NS4A-NS3 CO 865.5 84.5 83.5 84.0 650 20 AAX80362 HCV NS4A-NS3 CO 866 19 AAX22259 HCV NS4A-NS3	18		່ ທ	651	50	AAX80344	NS4A-NS3
20 863.5 85.0 1998 20 AAX80352 HCV NS4A NS3 CO 21 863.5 85.0 20198 20 AAX80356 HCV NS4A NS3 CO 22 861.5 84.8 651 20 AAX80356 HCV NS4A NS3 CO 23 860.5 84.7 1998 20 AAX80356 HCV NS4A NS3 CO 24 860.5 84.7 1998 20 AAX80356 HCV NS4A NS3 CO 25 86.7 2016 20 AAX80365 HCV NS4A NS3 CO 26 86.5 84.7 1998 20 AAX80365 HCV NS4A NS3 CO 27 868 84.6 648 20 AAX80365 HCV NS4A NS3 CO 28 857.5 84.4 651 20 AAX80363 HCV NS4A NS3 CO 29 857.5 84.4 650 20 AAX80363 HCV NS4A NS3 CO 29 857.5 84.4 650 20 AAX80362 HCV NS4A NS3 CO 29 857.5 84.4 650 20 AAX80362 HCV NS4A NS3 CO 29 857.5 84.6 650 20 AAX80362 HCV NS4A NS3 CO 31 853.5 84.0 651 20 AAX80360 HCV NS4A NS3 CO 32 853.5 84.0 651 20 AAX80350 HCV NS4A NS3 CO 34 85.1 86.0 651 20 AAX80350 HCV NS4A NS3 CO 36 86.5 83.5 1938 20 AAX32559 HCV NS4A NS3 CO 36 86.5 83.5 12980 19 AAX32559 HCV NS4A NS3 CO 37 86.5 83.5 12980 19 AAX3252 CODNA encoding hCV NS4A NS3 CO 41 848.5 83.5 12980 19 AAX3251 CODNA encoding hCV NS4A NS3 CO 42 844.5 83.1 7310 10 AAN92097 HCPAILLIS C VIT 43 844.5 83.1 7310 10 AAN92037 COMBINED OPEN T 44 844.5 83.1 7310 10 AAN92037 COMBINED OPEN T 45 844.5 83.1 7310 10 AAN92037 COMBINED OPEN T 46 844.5 83.1 7310 10 AAN93336 COMBINED OPEN T 47 844.5 83.1 7310 10 AAN93336 COMBINED OPEN T 48 844.5 83.1 7310 10 AAN93336 COMBINED OPEN T 48 844.5 83.1 7310 10 AAN93336 COMBINED OPEN T 48 AAA73334 STANDARTS 48 AAA73334 STANDARTS 48 AAA73334 STANDARTS 49 AAA73334 STANDARTS 40 HCPAILLIS C VITUS NS4A-NS3 fusion protease coding sequence #7.	19	÷.	Š.	99	20	AAX80357	NS4A-NS3
10 10 10 10 10 10 10 10	50	٠. د	٠. س	66	50	AAX80352	NS4A-NS3
23 860.5 84.7 651 20 AAX80343 HCV NS4A-NS3 CO 24 860.5 84.7 1958 20 AAX80356 HCV NS4A-NS3 CO 25 860.5 84.7 1998 20 AAX80356 HCV NS4A-NS3 CO 26 860 84.6 648 20 AAX80361 HCV NS4A-NS3 CO 27 860 84.4 650 20 AAX80351 HCV NS4A-NS3 CO 28 857.5 84.4 650 20 AAX80351 HCV NS4A-NS3 CO 30 857.5 84.4 650 20 AAX80351 HCV NS4A-NS3 CO 31 854 84.3 651 20 AAX80352 HCV NS4A-NS3 CO 32 853.5 84.0 650 20 AAX8036 HCV NS4A-NS3 CO 33 853.5 84.0 650 20 AAX8036 HCV NS4A-NS3 CO 34 861 85 84.0 650 20 AAX80350 HCV NS4A-NS3 CO 35 863.5 84.0 651 20 AAX80350 HCV NS4A-NS3 CO 36 848.5 83.5 12980 19 AAX53259 HCV NS4A-NS3 CO 37 848.5 83.5 12980 19 AAX5364 Hepatitis C vir 38 848.5 83.5 12980 19 AAX5364 Hepatitis C vir 40 848.5 83.5 12980 19 AAX5363 COmbined open 41 844.5 83.1 5300 10 AAN92003 COmbined open 42 844.5 83.1 7310 10 AAN92103 COmbined open 44 844.5 83.1 7310 10 AAN92103 COmbined open 45 844.5 83.1 7310 10 AAN92103 COmbined open 46 844.5 83.1 7310 10 AAN92103 COmbined open 47 844.5 83.1 7310 10 AAN92103 COmbined open 48 AA73334 standard; DNA; 594 BP.	22	. ر	U 4	7.5	207	AAX80360 AAX80348	NS4A-NS3
24 860.5 84.7 2016 20 AAX80356 HCW NS4A-NS3 CO	23	io	4	65	20	AAX80343	NS4A-NS3
25 800.5 84.7 2016 20 AAX80355 HCV NS4A-NS3 CO 26 860 84.6 648 20 AAX80355 HCV NS4A-NS3 CO 27 865 84.4 658 20 AAX80355 HCV NS4A-NS3 CO 28 857.5 84.4 651 20 AAX80351 HCV NS4A-NS3 CO 30 856.5 84.3 651 20 AAX80351 HCV NS4A-NS3 CO 31 855.5 84.4 651 20 AAX8035 HCV NS4A-NS3 CO 32 853.5 84.0 650 20 AAX8036 HCV NS4A-NS3 CO 33 853.5 84.0 650 20 AAX80350 HCV NS4A-NS3 CO 34 853.5 84.0 650 20 AAX80350 HCV NS4A-NS3 CO 35 848.5 84.0 651 20 AAX80350 HCV NS4A-NS3 CO 36 848.5 83.5 84.0 651 20 AAX80350 HCV NS4A-NS3 CO 37 848.5 83.5 1298.0 19 AAX23259 HCPATILIS C VIT 38 848.5 83.5 1298.0 24 ABX87286 HCPATILIS C VIT 41 844.5 83.1 5300 10 AAN92037 HCPATILIS C VIT 42 844.5 83.1 7310 10 AAN92103 COmbined open r 43 844.5 83.1 7310 10 AAN92103 COmbined open r 44 844.5 83.1 7310 10 AAN92103 COmbined open r 45 844.5 83.1 7310 10 AAN92103 COmbined open r 46 844.5 83.1 7310 10 AAN92103 COmbined open r 47 844.5 83.1 7310 10 AAN92103 COmbined open r 48 844.5 83.1 7310 10 AAN92103 COmbined open r 49 844.5 83.1 7310 10 AAN92103 COmbined open r 40 844.5 83.1 7310 10 AAN92103 COmbined open r 41 844.5 83.1 7310 10 AAN92103 COmbined open r 42 844.5 83.1 7310 10 AAN92103 COmbined open r 43 844.5 83.1 7310 10 AAN92103 COmbined open r 44 844.5 83.1 7310 10 AAN92103 COmbined open r 45 844.5 83.1 7310 10 AAN92103 COmbined open r 46 844.5 83.1 7310 10 AAN92103 COMBINED COMB	24	ö	÷.	66	50	AAX80356	NS4A-NS3
28 857.5 84.4 648 20 AAX80363 HCV NS4A-NS3 CO 29 857.5 84.4 655 20 AAX80363 HCV NS4A-NS3 CO 29 857.5 84.4 655 20 AAX80351 HCV NS4A-NS3 CO 30 855.5 84.4 651 20 AAX80342 HCV NS4A-NS3 CO 31 857.5 84.4 651 20 AAX80342 HCV NS4A-NS3 CO 32 857.5 84.0 651 20 AAX80342 HCV NS4A-NS3 CO 33 853.5 84.0 650 20 AAX80346 HCV NS4A-NS3 CO 34 853.5 84.0 650 20 AAX80350 HCV NS4A-NS3 CO 35 848.5 83.8 8145 20 AAX3259 HCV NS4A-NS3 CO 36 848.5 83.5 9646 24 AAX3259 HCV NS4A-NS3 CO 37 848.5 83.5 12980 19 AAV59361 HCPATILIS C VIT 38 848.5 83.5 12980 19 AAV59361 HCPATILIS C VIT 40 848.5 83.5 12980 19 AAV59361 HCPATILIS C VIT 41 844.5 83.1 5300 10 AAN90327 HCPATILIS C VIT 42 844.5 83.1 7310 10 AAN90327 COmbined open r 44 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 45 844.5 83.1 7310 10 AAN90336 COmbined open r 46 847.3 844.5 83.1 7310 10 AAN90336 COMBINED	25	٠ پ	• 4	7	200	AAX80361 AAX80365	NS4A-NS3
28 857.5 84.4 650 20 AAX80347 HCV NS4A-NS3 CO 30 856.5 84.4 651 20 AAX80351 HCV NS4A-NS3 CO 30 856.5 84.4 651 20 AAX80351 HCV NS4A-NS3 CO 31 856.5 84.3 651 20 AAX80352 HCV NS4A-NS3 CO 31 855.5 84.0 651 20 AAX80352 HCV NS4A-NS3 CO 33 853.5 84.0 651 20 AAX80346 HCV NS4A-NS3 CO 34 851 853.6 84.0 651 20 AAX80359 HCV NS4A-NS3 CO 34 851 83.8 8145 20 AAX23259 HCV NS4A-NS3 CO 34 848.5 83.5 9646 24 AAX80356 HCV NS3 DNA. H H PARTITIS C VIT AAX8035 HCV NS NSA CO 39 848.5 83.5 12980 19 AAV59361 HCPARTITIS C VIT AAX8025 HCV NS	27	വ	. 4	648	50	AAX80363	NS4A-NS3
29 857.5 84.4 651 20 AAX80351 HCV NS4A-NS3 CO 31 857.5 84.4 651 20 AAX80352 HCV NS4A-NS3 CO 32 853.5 84.0 650 20 AAX80350 HCV NS4A-NS3 CO 33 853.5 84.0 651 20 AAX80350 HCV NS4A-NS3 CO 34 851 83.8 8145 20 AAX23259 HCV NS3 DNA. H 35 848.5 83.6 1933 20 AAX23259 HCV NS3 DNA. H 36 848.5 83.5 12980 19 AAV59361 HCPATITIS C VIT 40 848.5 83.5 12980 19 AAV59361 HCPATITIS C VIT 40 848.5 83.5 12980 19 AAV59361 HCPATITIS C VIT 41 844.5 83.1 5360 10 AAN92097 HCPATITIS C VIT 42 844.5 83.1 7310 10 AAN92103 COmbined open r 44 844.5 83.1 7310 10 AAN92103 COmbined open r 45 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 46 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 46 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 47 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 46 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 47 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 48 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 49 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 40 844.5 83.1 7310 10 AAN92103 COMBINED OPEN R 40 848.5 83.1 7310 10 AAN92103 COMBINED OPEN R 40 848.5 83.6 83.1 7310 10 AAN92103 COMBINED OPEN R 40 848.5 83.6 83.1 7310 10 AAN92103 COMBINED OPEN R 40 848.5 83.1 7310 10 AAN	28	57.	4	650	20	AAX80347	NS4A-NS3
32 853.5 84.0 650 20 AAX80362 HCV NS4A-NS3 CO 33 853.5 84.0 651 20 AAX8036 HCV NS4A-NS3 CO 34 853.5 84.0 651 20 AAX20259 HCV NS4A-NS3 CO 35 848.5 84.0 651 20 AAX20259 HCV NS4A-NS3 CO 36 848.5 83.6 1933 20 AAX20259 HCV NS3 DNA. H 36 848.5 83.5 12980 19 AAV59361 HCPATITIS C VIR. 40 848.5 83.5 12980 19 AAV59361 HCPATITIS C VIR. 40 848.5 83.5 12980 24 ABX87285 HCPATITIS C VIR. 40 848.5 83.5 12980 24 AAX30207 HCPATITIS C VIR. 41 844.5 83.1 5360 10 AAV92037 HCPATITIS C VIR. 42 844.5 83.1 7310 10 AAN92103 Combined open r 44 844.5 83.1 7310 10 AAN92103 Combined open r 45 844.5 83.1 7310 10 AAN92103 Combined open r 45 844.5 83.1 7310 10 AAN92103 Combined open r 45 844.5 83.1 7310 10 AAN92103 HCPATITIS C VIR. AAA73334 AAA73334 Standard; DNA; 594 BP.	50 0	57.	÷ <	651	200	AAX80351	NS4A-NS3
32 853.5 84.0 650 20 AAX80346 HCV NS4A-NS3 co 34 853.5 84.0 651 20 AAX20259 HCV NS4A-NS3 co 34 853.5 84.0 651 20 AAX20259 Plasmid pET-BS(35 849 83.6 1933 20 AAX20259 Plasmid pET-BS(36 848.5 83.5 1980 19 AAV59361 Hepatitis C Vir. 37 848.5 83.5 12980 19 AAV59364 Hepatitis C Vir. 38 848.5 83.5 12980 24 ABX80285 Hepatitis C Vir. 40 848.5 83.5 12980 24 AAX80285 Combined open r. 41 844.5 83.1 5360 10 AAV92037 Hepatitis C Vir. 42 844.5 83.1 7310 10 AAV92103 Combined open r. 44 844.5 83.1 7310 10 AAV92103 Combined open r. 45 844.5 83.1 7310 10 AAV92103 Combined open r. 45 844.5 83.1 7310 10 AAV92103 Combined open r. 45 844.5 83.1 7310 10 AAV90336 Combined open r. 45 844.5 83.1 7310 10 AAV90336 Combined open r. 46 843.3 4 Standard; DNA; 594 BP.	2 14	95	• 4	648	20	AAX80362	NS4A-NS3
33 853.5 HCV NS4A-NS3 CO 34 851 83.6 HC 651 20 AAX20259 Plasmid pET-BS(35 848 83.6 1933 20 AAX20258 HCV NS3 DNA. H 36 848.5 83.5 1986 24 AAX99361 CONA encoding h 37 848.5 83.5 12980 24 ABK87285 HCP NS3 DNA. H 38 848.5 83.5 12980 24 ABK87285 HCP NCICOLING SQUAR 41 844.5 83.1 5300 10 AAV92097 CONDined open x 42 844.5 83.1 5300 10 AAN92097 CONDined open x 44 844.5 83.1 7310 10 AAN92103 CONDined open x 45 844.5 83.1 7310 10 AAN90327 CONDined open x 46 844.5 83.1 7310 10 AAN9033 CONDined open x 47 844.5 83.1 7310 10 AAN9033 CONDined open x 48 844.5 83.1 7310 10 AAN90336 CONDINED OPEN X 49 844.5 83.1 7310 10 AAN90336 CONDINED OPEN X 40 846.5 83.1 7310 10 AAN90336 CONDINED OPEN X 40 846.5 83.1 7310 10 AAN90336 CONDINED OPEN X 40 846.	32	53.	4	650	20	AAX80346	NS4A-NS3
35 82.0 AAX22239 FLASRING DEFTERS, 82.1 B49 83.6 8148.5 83.5 844.8 83.6 848.5 83.5 9446 19 AAV59361 Hepatitis C vir. 848.5 83.5 9446 19 AAV59361 Hepatitis C vir. 848.5 83.5 12980 19 AAV59364 Hepatitis C vir. 848.5 83.5 12980 19 AAV59364 Hepatitis C vir. 848.5 83.5 12980 19 AAV59364 Hepatitis C vir. 940 848.5 83.5 12980 19 AAV59364 Hepatitis C vir. 940 848.5 83.5 12980 19 AAV59364 Hepatitis C vir. 940 848.5 83.5 12980 19 AAV59364 Hepatitis C vir. 940 844.5 83.1 5360 10 AAV993097 Combined open r 42 844.5 83.1 5360 10 AAV993097 Combined open r 44 844.5 83.1 7310 10 AAV993106 Combined open r 45 844.5 83.1 7310 10 AAV99336 Combined open r 45 844.5 83.1 7310 10 AAV99336 Composite hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.	33	53.	٠,	9,	20	AAX80350	NS4A-NS3 CO
36 848.5 83.5 9646 24 ABK87285 Hepatitis C vir. 88 848.5 83.5 1980 19 AAV59361 Hepatitis C vir. 88 848.5 83.5 1980 19 AAV59364 Hepatitis C vir. 93 848.5 83.5 1980 19 AAV59364 Hepatitis C vir. 94 848.5 83.5 12980 19 AAV59364 Hepatitis C vir. 94 848.5 83.5 12980 24 ABK87286 Hepatitis C vir. 94 844.5 83.1 5300 10 AAN92097 Combined open r 44 844.5 83.1 5300 10 AAN92037 Combined open r 94 844.5 83.1 7310 10 AAN92103 Combined open r Composite hepatitis C vir. 1 AAA73334 AAA73334 Standard; DNA; 594 BP. Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.	ا م د	∩≪	٠, ۱4	40	200	AAX23259	asmid per-Ba
37 848.5 83.5 9646 24 ABK87285 CDNA encoding has 848.5 83.5 19980 19 AAV59364 Hepatitis C vir 40 848.5 83.5 12980 19 AAV59364 Hepatitis C vir 40 848.5 83.5 16622 21 AAZ36212 Combined open r 41 844.5 83.1 5300 10 AAN923297 Combined open r 43 844.5 83.1 5300 10 AAN92103 Combined open r 44 844.5 83.1 7310 10 AAN92103 Combined open r 45 844.5 83.1 7310 10 AAN92106 Composite hepat r AAA73334 AAA73334 standard; DNA; 594 BP. AAA73334 AAA73334 standard; DNA; 594 BP. Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.	9 6		,	4.4	9 6	AAV59361	ع کے
38 848.5 83.5 12980 19 AAV59364 Hepatitis C vir 40 848.5 83.5 12980 21 AAX58286 Hepatitis C vir 40 848.5 83.5 12980 21 AAX38212 Nucleotide sequing 41 844.5 83.1 5360 10 AAN92097 Combined open r 42 844.5 83.1 5360 10 AAN92103 Combined open r 44 844.5 83.1 7310 10 AAN92103 Combined open r 45 844.5 83.1 7310 10 AAN92103 Combined open r 45 844.5 83.1 7310 10 AAN92106 Composite hepat r 45 844.5 83.1 7310 10 AAN92106 Composite hepat r 45 844.5 83.1 7310 10 AAN92106 Composite hepat r 45 844.5 83.1 7310 10 AAN92103 Composite hepat r 45 844.5 83.1 7310 10 AAN93336 Composite hepat r 573334 AAA73334 standard; DNA; 594 BP.	37	48	, w	4	24	ABK87285	encoding
39 488.5 83.5 12980 24 AABK87286 Hepatitis C vir 410 844.5 83.5 16622 21 AACS6122 Combined open very 844.5 83.1 5360 10 AAN90207 Combined open very 844.5 83.1 5360 10 AAN90207 Combined open very 844.5 83.1 7310 10 AAN90203 Combined open very 844.5 83.1 7310 10 AAN90203 Combined open very 844.5 83.1 7310 10 AAN90336 Combined open very 845.5 844.5 83.1 7310 10 AAN90336 Combined open very 845.	38	₩.	ω.	298	19	AAV59364	citis C
41 844.5 83.1 5302 10 AAN92037 Conditioned open respectively. S 83.1 5360 10 AAN92037 Conditioned open respectively. S 83.1 5360 10 AAN92037 Combined open respectively. S 83.1 7310 10 AAN92103 Combined open respectively. S 83.1 7310 10 AAN92106 Combined open respectively. S 83.1 7310 10 AAN92106 Composite hepat AAA3334 standard; DNA; 594 BP. AAA73334 Standard; DNA; 594 BP. AAA73334; 19-DEC-2000 (first entry) Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.	66.4	6.0	m, u	298	2.5	ABK87286	titis C virus
42 844.5 83.1 5360 10 AAN90327 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.	7 7		, ,	2 6	7 -	AAL30212 BAN02002	tocide sed
43 844.5 83.1 6905 10 AAN92103 Combined open r 44 844.5 83.1 7310 10 AAN92106 Combined open r 45 844.5 83.1 7310 10 AAN90336 Combined open r Composite hepat r ALIGNMENTS ALIGNMENTS AAA73334 AAA73334 AAA73334 19-DBC-2000 (first entry) Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.	4.2	44	'n	36	101	AAN90327	open tea s C virus
44 844.5 83.1 7310 10 AAN92106 Combined open r 45 844.5 83.1 7310 10 AAN90336 Composite hepat ALIGNMENTS ALIGNMENTS AAA73334 AAA73334 AAA73334 AAA73334 Hepatitis C virus NS4A-NS3 fusion protease coding sequence #7.	43	4	ω.	6	10	AAN92103	oben r
ALIGNMENTS SULT 1 A73334 AAA73334 standard; DNA; 594 BP. AAA73334; 19-DEC-2000 (first entry) Hepatitis C virus NS4A-NS3 fusion protease coding sequence	4 4 5 5	ये ये	. m	33	10	AAN92106 AAN90336	ombined open r omposite hepat
A73334 AAA73334 standard; DNA; 594 BP. AAA73334; 19-DBC-2000 (first entry) Hepatitis C virus NS4A-NS3 fusion protease coding sequence							
AAA73334; 19-DBC-2000 (first entry) Hepatitis C virus NS4A-NS3 fusion protease coding sequence	ച് ന	73334	tandard	; DNA;	o h	BP,	
19-DBC-2000 (first entry) Hepatitis C virus NS4A-NS3 fusion protease coding sequence		333					
19-DEC-2000 (first entry) Hepatitis C virus NS4A-NS3 fusion protease coding sequence			•				
Hepatitis C virus NS4A-NS3 fusion protease coding sequence	7		ť		<u>~</u>		
		patitis		NS4A-N	~	protease coding	
1017 C 1011 C 1014 C 101			MC 2	4		took to the total	

Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds. Hepatitis C virus. Synthetic.

Location/Qualifiers Key

301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360

a

```
οy
                                                                                                                                                                     셤
                         ò
                                                 q
                                                                                 ò
                                                                                                           Q
                                                                                                                                                                                                                                                                                                                                                                   The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both resential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101 GlnalaproGlnGlySerargSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla 80
                                                                                                                                                                                                                                                                       Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                /*tag= a
/product= "NS4A-NS3 fusion protein #7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 594 BP; 105 A; 192 C; 151 G; 146 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     594
197
0
0
0
                                                                                                                                                                                                    Shang Y, Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-22 (1-197) x AAA73334 (1-594)
                                                                                                                                                                                                                                                                                                                                                Claim 26; Fig 17; 66pp; English.
                                                                                                                                                                        8
                                                                                                                                                                      (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3.95e-86
1016.00
100.00%
100.00%
                                                                                                                                                                                                    Weinheimer S,
                                                                                                                                           990S-0115271.
                                                                                                                 06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                               P-PSDB; AAB15225
                                                        WO200040707-A1
                                                                                                                                                                                                      Wittekind M,
                                                                                                                                            08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                                                                                    13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
 CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-7 variant.
                                                                                                                           180
                                                                                           480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                             ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                           481 CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA
 LeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArgGlySer
                Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease coding sequence #6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /*tag* a
/product= "NS4A-NS3 fusion protein #6"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 26; Fig 16; 66pp; English.
                                                                                                                                                                                                                                                                                           BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0115271
                                                                                                                                                                                                                                                                                           AAA73333 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                         (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2000-465976/40.
P-PSDB; AAB15224.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                        19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Synthetic.
                                361
                                                                                                                           161
                                                                                                                                                                                                                                                                                                                      AAA73333;
121
```

```
LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC 360
                                                                                  2000-465976/40.
                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                         P-PSDB; AAB15223
      WO200040707-A1
                                            38-JAN-1999;
                                                                                                                                                                                                                                                   Alignment Scores:
                   13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                   171
                                                                                                                                                                                                                                                                                                                                                                       41
                                                                                                                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  301
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                           81
                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                                                                                                                                                         Pred. No.:
                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                           a
ò
                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                            140
                                                                                                                                                                            240
                                                                                                                                                                                        100
                                                                                                                                                                                                     300
                                                                                                                                                                                                                               360
                                                                                                                                                                                                                                                         420
                                                                                                                                                                                                                                                                      160
                                                                                                                                                                                                                                                                                   480
                                                                                                                                                                                                                                                                                               180
                                                                                                                                                                                                                  GlnalaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                         120
                                                                                                                                     9
                                                                                                                                                               80
                                                                                   20
                                                                                               9
                                                                                                            40
                                                                                                                                     SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                           LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                  CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                   MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                               ATGAAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                           TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                    tis; NS3 protease; viral replication; chronic liver disease; fallure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A-NS3 fusion protease coding sequence #5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /*tag- a
/product- *NS4A.NS3 fusion protein #5*
      other;
                         594
0
0
0
0
0
                        Length:
Matches:
Conservative:
Mismatches:
     0
     ..
⊢
                                                  Indels:
Gaps:
     147
                                                                     (1-594)
     .;
3
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
     152
     <del>ن</del>
                                                                     US-09-965-594-22 (1-197) x AAA73333
                                                                                                                                                                                                                                                                                                                                                                BP
                        1.44e-85
1010.00
99.49%
99.49%
99.41%
     BP; 104 A; 191
                                                                                                                                                                                                                                                                                                                                                                DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                         19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
Synthetic.
                                            Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis; NS3
                                      Percent Similarity:
     Sequence 594
                   Alignment Scores
                                                                                                            21
                                                                                                                                                  121
                                                                                                                                                                                                                                                                                                           481
                                                                                                                                                                                                                                                                                                                                                                AAA73332
                                                                                                                                      41
                                                                                                                                                                            181
                                                                                                                                                                                                     241
                                                                                                                                                                                                                                                         361
                                                                                                                                                                                                                                                                      141
                                                                                                                                                                                                                                                                                  421
                                                                                                                                                                                                                                                                                               161
                                                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                  101
                                                                                                                                                                                                                                301
                                                                                                                                                                                                                                            121
                                                                                                                                                               61
                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                           AAA73332
                                                 Query Match:
DB:
                          :.
₽
                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                              g
                                                                                                           ô
                                                                                                                                                                          g
                                                                                                                                                                                        ò
                                                                                                                                                                                                   a
                                                                                                                                                                                                                ò
                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                     셤
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                      Q
                                                                                                                                  ò
                                                                                                                                                 a
                                                                                                                                                               ò
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TyralaClnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlnalaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAAAACCAGGTTGAAGGTGAAGTTCAGATGGTTTCCACCGCTACCCAGACCTTCCTGGCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 594 BP; 105 A; 189 C; 153 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        594
1193
2
0
0
                                                                                                                                                              Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x AAA73332 (1-594)
                                                                                                                                                           Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                     Claim 26; F1g 15; 66pp; English.
                                                                                                           8
                                                                                                        BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3.64e-84
995.00
98.98%
97.97%
97.93%
06-JAN-2000; 2000WO-US00345
                                                  99US-0115271
                                                                                                                                                              Wittekind M, Weinheimer S,
```

9

```
Alignment Scores:
                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A procease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and Niver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                 540
                                                                                                                           480
                                                                                                                                                                              180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                           ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                              481 CCGCCTGTCACGCTGTTGGTATCTTCCGTGCTGTTTCCACCCGTGGTGTTGCTAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus NS4A-NS3 fusion protease coding sequence #4.
                                                                                                                                                                                                                                                                                                            /*tag= a
/product= "NS4A-NS3 fusion protein #4"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       594 BP; 105 A; 187 C; 155 G; 147 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Υ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2hang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 26; Fig 14; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                          BP
                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73331 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2000-465976/40.
P-PSDB; AAB15222.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-JUL-2000
                                                                                                                                                                              161
                                                                                                                                                                                                                                                                                 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                        361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73331;
                                                                                                                                                                                                                                                                                                                                                                                  AAA73331
AAA7331
AAA731
AAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Key
                                                                                                                                                                                                                                                                                                                    셤
                                                                           ò
                                                                                                                      g
                                                                                                                                                              ŏ
                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                 ò
```

```
160
                                                                                                                                                                                                                                                                                                                                                                                     ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                 240
                                                                                                                                                                                                                                    100
                                                                                                                                                                                                                                                                          120
                                                                                                                                                                                                                                                                                          360
                                                                                                                                                                                                                                                                                                             140
                                                                                                                                                                                                                                                                                                                                420
                                                                                                                                                                                                                                                                                                                                                                   480
                                                                                                                                                                                                                                                                                                                                                                                                       481 CCGCTCGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA 540
                                                                                                                                                                              180
                                                                                                                                                                                                                                                     300
                                                                                                                                                             9
                                                                                                                                                                                                80
                                                                                                                       40
                                                                                                     9
                                                                                  301 CAGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                  LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLys
                                                                                                                       21 TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                             LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                             LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease coding sequence #3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /*tag= a
/product= "NS4A-NS3 fusion protein #3"
594
190
2
0
0
                   Conservative:
Mismatches:
           Matches:
Length:
                                      Indels:
Gaps:
                                                                US-09-965-594-22 (1-197) x AAA73331 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ВР
9.23e-83
980.00
97.46%
96.45%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAA73330 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus.
Synthetic.
                Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-DEC-2000
                                                                                                                                                            41
                                                                                                                                                                              121
                                                                                                                                                                                                                                                                                                                                                                                     161
                                                                                                                                                                                                                                                                                                                                                 141
                                                                                                                                                                                                                                                                         101
                                                                                                                                                                                                                                                                                                                                361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAA73330;
                                                                                                                                                                                                                                                                                                             121
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 5
```

5

```
Synthetic
                                                                                                                                                                            AAA73335;
                                                                                                                            RESULT
                                                                                                                                                    à
                                             q
                                                                       ò
                                                                                             a
                                                                                                                                                                                                                          The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the Polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutates and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TACGCTCAGCAGACTCGAGGTGAGGAGGGTTGCCAAGAAACCTCCCAGACCGGTCGTGAC 120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TyralacinchnthrargciyclucinciythrcinrysthrserHisThrciyArgasp 40
                                                                                                                                        Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BP; 103 A; 186 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                    594
187
200
00
                                                                          Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                           protease and protease inhibitor complexes. The sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (1-594)
                                                                          Zhang Y,
                                                                                                                                                                                                     26; Fig 13; 66pp; English.
                                                 ဗ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-22 (1-197) x AAA73330
                                                 (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                    3.59e-81
963.00
95.94%
94.92%
                                                                          Weinheimer S,
06-JAN-2000; 2000WO-US00345
                         99US-0115271
                                                                                                   2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                  P-PSDB; AAB15221
                                                                                                                                                                                                                                                                                                                                                                                               Sequence 594
                                                                           Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                          Alignment Scores:
                         08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       241
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                       Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                        ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
```

g g

g ă

ò

ò ď οŽ à

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this
                                                                                        180
                                                                                                                                              540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis {\sf C}
161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                      NS3 protease; viral replication; chronic liver disease; ure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                        181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus NS4A-NS3 fusion protease coding sequence #8.
                                                                                                                                                                                                                                          sequence contains the alpha-helix0 wild-type sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /product- "NS4A-NS3 fusion protein #8"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 594 BP; 98 A; 189 C; 153 G; 154 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2hang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
1..594
/*tag= a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Disclosure; Fig 18; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8
                                                                                                                                                                                                                                                                                                                                                                                  ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1.21e-78
936.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0115271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                  AAA73335 standard; DNA; 594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   P-PSDB; AAB15226.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis;
```

```
Score:
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
 ò
                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           à
                                                                                                                                                                                                                                                                                                       å
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                     160
                                                                                                                                                                  100
                                                                                                                                                                               300
                                                                                                                                                                                             101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                           360
                                                                                                                                                                                                                         140
                                                                                                                                                                                                                                480
                                                                                                                                                                                                                                                                                  180
                                                                                                                                                                                                                                                                                               540
                                                                                                                       180
                                                                                                                                                    240
                                                  50
                                                                9
                                                                              40
                                                                                                          09
                                                                                                                                      80
                                                                                                                                                                        LeuLeuSerProArgPro11eSerTyrLeuLysG1ySerSerG1yG1yProLeuLeuCys
                                                                                                                                                                                                                                                                                        SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                   LeuvalThrArgHisAlaAspvalIleProvalArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                   CTGCTGTCCCCGCGTCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCGCTGCTGTGC
                                                                                                                                                                                                                                                                                 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                      ThrSerIleasnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                    181 ACCIGCATCAACGGIGITIGCIGGACCGITIACCACGGIGCIGGIACCCGIACCAICGCI
                                                                                                          LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                 MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                         TyralaGlnGlnThrargGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                              ltis; NS3 protease; viral replication; chronic liver disease;
failure; liver cancer; mutant; mutein; ds.
                                                                                                                                                                                                                                                                                                             Hepatitis C virus NS4A-NS3 fusion protease coding sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "NS4A-NS3 fusion protein #2"
0050
 Conservative:
       Mismatches:
Indels:
                                    US-09-965-594-22 (1-197) x AAA73335 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                       ВP
                                                                                                                                                                                                                                                                                                                                                       AAA73329 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product= '
                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
93.91%
93.91%
92.13%
                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus. Synthetic.
       Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis; NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200040707-A1
Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                  19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13-JUL-2000
                                                                                                          41
                                                                                                                                      61
                                                                                                                                                                               241
                                                                                                                                                                                                          301
                                                                                                                                                                                                                                                     141
                                                                                                                                                                                                                                                                  121
                                                                                                                                                                                                                                                                                 161
                                                                                                                                                                                                                                                                                              481
                                                                                                                                                                                                                                                                                                             181
                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                     AAA73329
               Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                      liver
                                                                                                                                                                                                                                                                                                                                        Key
                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                            g
                                                                                                                                                                                                         g
                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                             à
                                                                                                                                                                                                                                                                                                                           g
                                                  ö
                                                              g
                                                                             ò
                                                                                          g
                                                                                                        ò
                                                                                                                     q
                                                                                                                                     δ
                                                                                                                                                 g
                                                                                                                                                                ŏ
                                                                                                                                                                                           οŽ
                                                                                                                                                                                                                                                                                ογ
```

```
The present sequence is the coding sequence for a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         294
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       141 LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9
                                                                                                                                                                           Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21 TyralaglnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41 LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  121 LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              588
183
29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                              >
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches:
Conservative:
Mismatches:
                                                                                Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-965-594-22 (1-197) x AAA73329 (1-588)
                                                                              Zhang Y,
                                          ဗ
                                                                                                                                                                                                                                                                              Claim 26; Fig 12; 66pp; English.
                                        (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5.38e-78
929.00
94.42%
92.89%
91.44%
                                                                              Weinheimer S,
99US-0115271
                                                                                                                    2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                         P-PSDB; AAB15220
                                                                              Wittekind M,
08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   81
```

40

20 54 174

80

9

234

g

ð

qq

ò

```
101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                            195 CCGGCTCCGCAGGGTTCCCGTTCCCTGACCCCGTGCACCTGCGGGTTCCTCCGACCTGTAC 354
                                                                                                                                                                                                                                                                                                                                                                                                       LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                                                                                                                                                                                                                                                               LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer
                                                                                                                                                MetLysLysLysClySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                             55 TACGCTCAGCAGACTCGAGGTCTGCTGCATCATCACCTCCCTGACCGGTCGTGAC
                                                                                       21 TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181 AlavalAspPheIleProvalGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pestivirus; Npro; protease; NS3; screening; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
US-09-965-594-22 (1-197) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ABA95615 standard; DNA; 12734 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           990S-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hong 2, Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                04-DEC-2001
                                                                                                                                                                              115
                                                                                                                                                                                                                                                                                                                                                                                                                                               141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             535
                                                                                                                                                                                                          61
                                                                                                                                                                                                                                     175
                                                                                                                                                                                                                                                                   81
                                                                                                                                                                                                                                                                                               235
                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                  355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ABA95615;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABA9561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                  ç
                                                                                                                                                                    q
                                                                                                                                                                                                š
                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                       셤
                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                    ద
                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ОD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                               ò
                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis. Ilver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
161 ProAlaClyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                 Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                 Hepatitis; NS3 protease; viral replication; chronic liver disease:
liver failure; liver cancer; ds.
                                                         Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      588
181
13
13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                  /product- "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
1..588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Fig 10; 66pp; English.
                                                                                                                                                   AAA73328 standard; DNA; 588 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1.81e-75
902.00
92.39%
91.88%
88.78%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271
                                                                                                                                                                                                          19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                      /*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          P-PSDB; AAB15212.
                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         06-JAN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               13-JUL-2000
                                                                                                                                                                              AAA73328;
```

Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV)

œ

AAX80355;

```
The present invention relates to a nucleic acid construct encoding a chimeric Hepatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimmeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                       532
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     593 TGCATCAATGGGGTATGCTGGACTGTCTACCACGGGGCCGGAACGAGGACCATCGCATCA 652
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               81
pestlvirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                         GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GCCCAGCAGACGAGAGGCCTCCTAGGGTGTAAGATCACCAGTCTGACTGGCCGGGACAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                 Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-22 (1-197) x ABA95615 (1-12734)
                                                               Example 2; Columns 17-28; 20pp; English.
                                                                                                                                                                                                                                                                                          5.54e-72
882.50
92.31%
90.26%
86.86%
                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  893
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                å
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

AAX80355 standard; cDNA; 1998 BP

953

RESULT 10 AAX80355 ID AAX8 XX

```
243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NS4A-NS3 complex comprising a central hydrophobic domain of native NCV NS4A Peptide, a linker, and an HCV NS4 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminum of the HCV NS4A peptide is tethered by the linker sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination of mode of binding of HCV inhibitors by NNR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the Arpase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention describes a covalent hepatitis C virus (HCV)
                                                                               HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1998
167
15
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 1998 BP; 411 A; 595 C; 569 G; 423 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative:
Mismatches:
Indels:
                                                       HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                            ż
                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Page 166-169; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                   New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x AAX80355 (1-1998)
                                                                                                                                                                                                                                                                                                                                            Weber PC,
                                                                                                                                                                                                                                                                     98US-0094331.
                                                                                                                                                                                                                                           98WO-0S24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2.5e-72
875.50
92.86%
85.20%
86.17%
                            (first entry)
                                                                                                                                                                                                                                                                                                                                            Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                 (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                     28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
Pred. No.:
                                                                                                                                                                                    WO9928482-A2
                                                                                                                                                                                                                                          24-NOV-1998;
                            07-SEP-1999
                                                                                                                                                                                                               10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           62
                                                                                                                                                         Synthetic
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ð
```

us-09-965-594-22.rng

other;

569 G; 423 T; 0

ö

BP; 411 A; 595

```
Sequence 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10-JUN-1999
                     Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAX80354;
                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                              424
                                                                                                                                                                                                                                                                                                                                                                                                                  142
                                                                                                                                                                                                                                                                                                                                                                                                                                       184
                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAX80354
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ο̈́
                                                                                                                                             g
                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                   8
                                                                                                                                                                                                                                                                                                                                                                          ογ
                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                          ò
                                                                                                                                                                  à
                                                                                                                                                                                    a
                                                                                                                                                                                                         ò
                                                                                                                                                                NS4A FNS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASE activity of NS3. The covalent Covalent covalent complexes are more soluble, stable and active than the non-
                                                                                             543
         121
                                                    122 ValThrargHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                          LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
 ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                     102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                                                                      HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                          ValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                     604 GIGGACTTIGGCCCGFAGAGTCCATGGAAACTACTATGCGGTCTCCG 651
                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Page 179-182; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PC,
                                                                                                                                                                                                                                                                       BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Weber
                                                                                                                                                                                                                                                                       AAX80359 standard; cDNA: 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98WO-US24528
                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28-JUL-1998;
                                                                                                                                                                                                                                                                                                                07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999
                                                                                                                           142
                                                                                                                                                                                                           182
                                                                                                                                                                                                                                                                                            AAX80359;
                                                                                                                                                                                                                                                    RESULT 11
                                                                                                                                                                                                                                                             a
                                                                                                                                                                                     qq
                 g
                                         ò
                                                        a
                                                                                  ò
                                                                                                                          ö
                                                                                                                                            a
                                                                                                                                                              õ
                                                                                                                                                                                                            õ
                                                                                                                                                                                                                         셤
```

```
243
                                                                                                                                                                                                                                                                                                    303
                                                                                                                                                                                                                                                                                                                              101
                                                                                                                                                                                                                                                                                                                                                      363
                                                                                                                                                                                                                                                                                                                                                                                121
                                                                                                                                                                                                                                                                                                                                                                                                        423
                                                                                                                                                                                                                                                                                                                                                                                                                                ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                         483
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     123
                                                                                                                                                                                               183
                                                                                                                                                                                                                         61
                                                                                                                                                                                                                                                                           81
                                                                                                                      21
                                                                                                                                                                        47
                                                                                                                                                                                                                                                                                                                           102 AlaproGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
||||||||
364 GCGCCCCCGGGGGGGCGTTCCTTGACACCATGCACCTGTGGCAGCTCAGACCTTTACTTG
                                                                                                                                                                      LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                             62 SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                      GlySerValValIleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1998
166
16
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:104
                            Conservative:
Mismatches:
Indels:
    Length:
Matches:
                                                                    Gaps:
                                                                                           (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAX80354 standard; cDNA; 1998 BP
                                                                                            US-09-965-594-22 (1-197) x AAX80359
4.77e-72
872.50
92.86%
84.69%
85.88%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus
Synthetic.
                                          Best Local Similarity:
                              Percent Similarity:
```

```
22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys 41
                                  Synthetic.
                                                                                                                            AAX80345;
                                                                                 RESULT 13
                                                                                          AAX80345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score:
                                                                                                                            염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  244 TGCGTCAACGGCGTGTGTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LeuSerProArgProIleSerTyrLeuLysClySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln 101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
                                                                                                                                                                              NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, when the helicase activity and the Arpase activity of NS3. The covalent was the helicase activity and the Arpase activity of NS3. The covalent was the model of the model of NS3. The covalent was the model of NS3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              81
                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlySerValValIleValGlyArgIleAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CTCTCCCCCAGGCCTGTCTCCTACTTGAAGGGCTCTTCGGGTGGTCCACTGCTCTGCCCT
                                                                                                                                                                                                                                                                                      NS4A-NS3 complexes are more soluble, stable and active than the non-
                                                                                                                                                                       present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                                 covalent protease-peptide complexes previously available
                                                                                                                                                                                                                                                                                                                       Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                       1998
167
14
12
3
                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                ż
                                                                                                                                                                                                                                                                                                                                                                                                      Indels:
                                                                                                                                              Disclosure; Page 163-166; 211pp; English.
                                                                                                                            New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-22 (1-197) x AAX80354 (1-1998)
                                                                                Weber PC,
                                                                                                                                                                                                                                                                                                                                                      5.91e-72
871.50
92.35$
85.20$
85.78$
98WO-US24528
                      98US-0094331
97US-0067315
                                                                                Malcolm BA, Taremi SS,
                                                        (SCHE ) SCHERING CORP
                                                                                                    WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                       Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                      28-JUL-1998;
28-NOV-1997;
24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      484
                                                                                                                                                                                                                                                                                                                                                                                                                                                           2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               304
                                                                                                                                                                                                                                                                                                                                                          Pred. No.:
à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    염
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ដ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
```

```
64 GGTTCTGTTGTTGTTGGTAGAATTATTTTATCTGGTAGTGGTAGTATATCACGGCCTAC 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5 GlySerValVallleValGlyArglleAsnLeuSerGlyAsp------ThrAlaTyr 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present invention describes a covalent hepatitis C virus (HCV) NS4A.NS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent NS4A NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                         HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 651 BP; 120 A; 187 C; 200 G; 144 T; 0 other;
                                                                                                                                                                                           HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:95
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; Page 147-148; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-965-594-22 (1-197) x AAX80345 (1-651)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weber PC,
   ВР
AAX80345 standard; cDNA; 651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                868.50
92.82%
85.13%
85.48%
                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                          07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10-JUN-1999.
```

```
ValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141
                                                                                                                                                                                                                                                           AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181
NS4A peptide, a linker, and an HCV NS3 serinc protease domain, where the hydrophobic domain of native HCV NS3 protease domain. The present sequence encodes an example of the above complex. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. The year also be used for detecting inhibitors by CAP protease activity, the helicase activity and the ATPASE activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
                                                                                                                                                                                                                                                                                                     102 AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu
                                                                                                                                                                                                                                                                                           22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                     AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                               SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                 GlySerValVallleValGlyArglleAsnLeuSerGlyAsp-----ThrAlaTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1998
166
15
12
3
                                                                                                                    Sequence 1998 BP; 410 A; 596 C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CDNA SEQ ID NO:103.
                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                            US-09-965-594-22 (1-197) x AAX80358 (1-1998)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          complex encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAX80353 standard; cDNA; 1998
                                                                                                                                                   1.13e-71
868.50
92.35%
84.69%
85.48%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                                                                                                                                 Best Local Similarity:
                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3
                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    07-SEP-1999
                                                                                                                                                                                                                                                                                                                                     42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               162
                                                                                                                                                                                                                                                                                                                                                                               62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               424
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             604
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX80353;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         484
                                                                                                                                                                                                                                                S
                                                                                                                                                                                                                                                                                                                                                                                                                                            304
                                                                                                                                                                                               Query Match:
                                                                                                                                                     ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 1:
                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          XEXEXEX
                                                                                                                                                                                                                                                                                                                                                                                                                                          đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
  qa
                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                          à
                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                           483
                                                                                                                                                                                                                                                161
                                                                                                                                                                                                                                                                    543
                                                                                                                                                                                                                                                                                         181
                                                                                                                                                                                                                                                                                                               603
  303
                                                                                                                    101
                                                                                                                                         363
                                                                                                                                                             121
                                                                                                                                                                                  423
                                                                                                                                                                                                        141
                                                     243
                                61
                                                                          81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV
                                                                                                                               162 AlaGlyHisAlaVaiGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla
                                            ProLysGlyProValThrGlnMctTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                               LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro
                                                                                                                                                                                                                                                          544 TCGGGGCACGCTGTGGGCATCTTCCGGGCTGCCGTATGCACCCGGGGGGTTGCGAAGGCG
                                                                                               TGCGTCAACGGCGTGTTGGACCGTTTACCATGGTGCTGGCTCAAAGACCTTAGCCGGC
                                                                                                                                                                                                       ValThrArgHisAlaAspValilcProValArgArgArgGlyAspSerArgGlySerLeu
                                                                           SerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer
                                AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                   ValAspPheIleProValGluSerLeuGluThrThrMetArgSer 196
                                                                                                                                                                                                                                                                                                                                               604 GTGGACTTTGTGCCCGTAGAGTCCATGGAAACTACTATGCGGTCT 648
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV NS4A-NS3 complex encoding cDNA SEQ ID NO:108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 176-179; 211pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wew hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PC,
                                                                                                                                                                                                                                                                                                                                                                                                    ВP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Weber
                                                                                                                                                                                                                                                                                                                                                                                                  AAX80358 standard; cDNA; 1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BA,
                                                                                                                                                                                                                                                                                                                                                                                                                                            07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
                                                                                                                                                                                                                                                                                                                                   182
                                                                          62
                                                                                               244
                                                                                                                    82
                                                                                                                                        304
                                                                                                                                                             102
                                                                                                                                                                                  364
                                                                                                                                                                                                                                                142
                                                                                                                                                                                                                                                                   484
                                                                                                                                                                                                                                                                                                                                                                                                                       AAX80358;
                                                                                                                                                                                                       122
                                                                                                                                                                                                                           424
           124
                                42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm
                                                                                                                                                                                                                                                                                                                                                                             셤
           a
                                δ
                                                   g
                                                                   ŏ
                                                                                         a
                                                                                                                   ò
                                                                                                                                   g
                                                                                                                                                           9 9
                                                                                                                                                                                                     ολ
                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                              ç
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                   ò
```

61

41

243

8.1

483

```
The present invention describes a covalent hepatitis C virus (HCV) NSAA-NS3 complex comprising a central hydrophobic domain of native HCV NSAA peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NSAA peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The present sequence encodes an example of the above complex. The covalent NSAA-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPASE activity of NS3. The covalent NSAA-NS3 complexes are more soluble, stable and active than the noncovalent protease-peptide complexes previously available.
HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; 568 G; 424 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                 Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Disclosure; Page 160-162; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                         New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                               Weber PC,
                                                                                                                                                                                                             98WO-US24528
                                                                                                                                                                                                                                                   98US-0094331
97US-0067315
                                                                                                                                                                                                                                                                                                                                               Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                         (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-385385/32
                                                                                 Hepatitis C virus
                                                                                                                                       W09928482-A2
                                                                                                                                                                                                                                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                             24-NOV-1998;
                                                                                                                                                                           10-JUN-1999
                                                                                                 Synthetic.
```

```
AlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeu 121
                                                                                                                                                          243
                                                                                                                                                                                                                                    22 AlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                     42 AsnGlnvalGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLcuAlaThr 61
                                                                                                                                                                                                                          62 SerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSer 81
                                                                                                                                                                                               1998
166
15
12
                          Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                         Gaps:
                                                                                          US-09-965-594-22 (1-197) x AAX80353 (1-1998)
Sequence 1998 BP; 410 A; 596
                          1.4e-71
867.50
92.35%
84.69%
85.38%
                                             Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                  Alignment Scores:
                                                                                                                                                                                                                                                                                                    102
                                                                                                                                                                                                                                                                                 a
                                                                                                                             q
                                                                                                                                               ò
                                                                                                                                                                 q
                                                                                                                                                                                                     qq
                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                    ò
                                                                                                             ò
                                                                                                                                                                                     ò
```

123	141	483	191	543	181	603			
364 GCGCCCCCCGGGGCGCGTTCCTTGACATCCATGCACCTGTGGCAGCTCAGACCTTTACTTG 423	ValThrargHisalaaspValIleProValArgArgArgGlyAspSerArgGlySerLeu 141	GTCACGAGACATGCTGACGTCATTCCGGTGCGCCGGGGGGGG	LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyBroLeuLeuCysPro 161	-0	162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLysAla 181	TCGGGGCACCTGTGGGCATCTTCCGGGCTGCTGTTTCTCCCGGGGGGTTGCGAAGGCG	ValAspPheileProValGluSerLeuGluThrThrMetArgSerPro 197	- 0	
364	122	424	142	4 84	162	544	182	604	
qq	QY	qq	ογ	qq	ογ	q	ογ	qa	

Search completed: August 30, 2003, 19:48:21 Job time : 188.939 secs

•

```
Sequence 23, Application US/09965594

Sequence 23, Application US/09965594

Patent No. US20020106642A1

GENERAL INFORMATION:
APPLICANT: Wittekind, Michael
APPLICANT: Welnhelmer, Steven
APPLICANT: Welnhelmer, Steven
APPLICANT: Chang, Yaden
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Modified Forms of Protease: Inhibitor Complexes
FILE REFERENCE: DB17Sequences
CURRENT FILING DATE: 1999-01-08
FRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 23
LENTH: 594
                                                                                                         Sequence 4, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
                                                                                                                                                                                                                                                                                                19, Appl
19, Appl
19, Appl
2, Appli
1, Appli
2, Appli
120, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                17, Appl
17, Appl
13, Appl
13, Appl
16, Appl
1, Appl
8, Appl
15, Appl
                                                                                                                                                                                                                                           Sequence 5, Sequence 5, Sequence 19 Sequence 19 Sequence 19 Sequence 2, Sequence 2, Sequence 2, Sequence 17 Sequence 17 Sequence 18 Sequence 16 Sequen
                                                        Sequence
Sequence
Sequence
Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence Sequence S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence Sequence Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                              US-09-238-076-19
US-09-238-076-19
US-09-965-594-19
US-09-965-594-2
US-09-881-239-2
US-09-921-397-120
US-09-921-397-120
US-09-921-397-120
US-09-921-397-120
US-09-921-397-120
US-09-929-955-13
US-09-919-901-1
US-09-919-901-1
US-09-919-901-15
US-09-919-901-15
US-09-919-901-15
US-09-884-456-69
US-10-259-275-41
US-10-005-469-1
US-09-965-594-21
US-09-965-594-19
US-09-965-594-17
US-09-965-594-15
US-09-965-594-13
US-09-965-594-13
US-09-965-594-13
US-09-965-594-4
US-10-133-133A-6
                                                                                                                                                                 US-09-238-076-1

US-09-995-937-1

US-09-917-58-1

US-09-238-076-5

US-09-995-937-5

US-09-917-563-5
   TYPE: DNA
CRCANISM: Hepatitis C virus
US-09-965-594-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9416
20416
13910
13910
13910
13910
13910
13910
7992
7992
7992
7992
7992
7992
8638
8638
8638
 5944
5944
5944
5944
612
612
9646
9646
9646
12980
12980
12980
9379
9416
                                                                                                                                                                                                                                                                                                                                                                                                             836
10803
10803
841.5
841.5
841.5
831.5
831.5
831.5
831.5
821.5
821.5
821.5
821.5
821.5
821.5
821.5
821.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
                                      "MODEL-frame._prin.model - DEV-xlp
-Q-/cgn2_1/USPTO_spool/US09965594/runat_29082003_151920_28367/app_query.fasta_1.2872
-Q-/cgn2_1/USPTO_spool/US09965594/runat_29082003_151920_28367/app_query.fasta_1.2872
-DB-vDb1shad.Applications_NA - OFFWT-fastap - SUFFIX-rnpb - WIRMATCH=0.1
-DOPPCI—0 - LOOPEXI—0 - UNITS—bits - START=1 - END - 1 - MATRIX-blosum62
-TRANS-human40.cdi - LIST-45 - DOCALIGN-200 - THR_SCORE-pot - THR_MXX-100
-THR_MIN-0 - ALIGN-15 - MODE-LOCAL - OUTFWT-pto - NORM-ext - HEAPSIZE-500 - MINLEN-0
-MAXLEN-2000000000 - USER-US09965594_eCGN_1.1_E864_ernaft_-DS082003_15192_28367
-MODUG - CICPG-3 - NO_MMAP - LARGEQUERY - NGC SCORES-0 - WAIRT_-DSPBLOCK-100
-LONGLOG - DEV_IIMEOOT-120 - WARN_TIMEOOT-30 - THREADS-1 - XGAPOP>-10 - XGAPEXI-0.5
-FGAPOP=6 - FGAPEXI-7 - YGAPOP-10 - YGAPEXI-0.5 - DELOP=6 - DELEXI-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 23, Appl
                                                                                                                        August 30, 2003, 19:26:03; Search time 176.482 Seconds (without alignments) 2560.981 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                         1 MKKKGSVVIVGRINLSGDTA.....VAKAVDFIPVESLETIMRSP 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /cgn2_6/ptodata/2/pubpna/USO7_PUBCOMB.Seq:*
/cgn2_6/ptodata/2/pubpna/DCT_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/DCT_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO6_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO6_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO8_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/USO9_NEW_PUB.seq:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
/cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
                 GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                           - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-23
                                                                                                                                                                                                                                                                                                                                                                                                                              Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                         1533700 segs, 1147125425 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Published_Applications_NA:*
                                                                                                                                                                                                                                                                                                 0.5
                                                                                                                                                                                                                                                                      BLOSUM62
Xgapop 10.0 , Xgapext (
Ygapop 10.0 , Ygapext (
Fgapop 6.0 , Fgapext nelop 6.0 , Delext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    594 10
                                                                                                                                                                                                      US-09-965-594-22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query
Match Length DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    100.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1016
                                                                                                                                                                                                                                                                                Scoring table:
                                                                                                                                                                                                                         Perfect score:
                                                                                             OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database :
                                                                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  <u>₹</u>
                                                                                                                                Run on:
```

Result

a

```
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
US-09-965-594-19
                 Alignment Scores
JS-09-965-594-21
                                                          Query Match:
DB:
                          ..
⊗
                                  Score:
                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                  οχ
                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                 g
                                                                                                                                   ò
                                                                                                                                                 qq
                                                                                                                                                                                g
                                                                                                                                                                                                                q
                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                      õ
                                                                                                   ö
                                                                                                                                                                   ^{\circ}_{Y}
                                                                                                                                                                                                   ô
                                                                                                                                                                                                                    120
                                                                                                                                                                                                    240
                                                                                                  9
                                                                                                                  40
                                                                                                                                                                                    80
                                                                                  50
                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 21, Application US/09965594

Fatent No. US20020106642A1

GENERAL INFORMATION:

APPLICANT: Wittekind:

APPLICANT: Wainheimer, Steven

APPLICANT: Wainheimer, Steven

APPLICANT: Coldfarb, Valentina

APPLICANT: Coldfarb, Valentina

TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studi

TITLE OF INVENTION: Facilitating Inhibitor Complexes

TITLE OF INVENTION: OF Protease: Inhibitor Complexes

FILE REFERENCE: DB17Sequences

CURRENT FILING DATE: 201-09-27

FRIOR APPLICATION NUMBER: 60/115,271

PRIOR PLICATION NOS: 26

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 21

**LENGTH: 594
                                                                                                                                                                                   MetLysLysLysGlyServalvalileValGlyArglleAsnLeuSerGlyAspThrAla
                                                                                                   1 ATGAAAAAAAAAAAGGATCCGTTGTTATCGTCGGCCGTATCAACCTGTCCGGTGACACCGCT
                                                                                                                  21 TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                       594
197
0
0
0
       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                  US-09-965-594-22 (1-197) x US-09-965-594-23 (1-594)
                                                   Gaps:
     4.57e-109
1016.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE: DNA ORGANISM: Hepatitis C virus
                     Percent Similarity:
Best Local Similarity:
Query Match:
DB:
Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 2
US-09-965-594-21
                                                                                                                                                                                                 a
                                                                                                                                                                                                                    07
                                                                                                                                                                                                                                                                                                   9 % g
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                               g
                                                                                                                              Oy
Oy
                                                                                                                                                                               ογ
                                                                                                                                                                                                                                                     ογ
                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                  å
                                                                                                                 δy
                                                                                                                                                              g
```

```
APPLICANT: Wittekind, Michael
APPLICANT: Wittekind, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Zhang, Yaqun
APPLICANT: Zhang, Yaqun
APPLICANT: Chang, Yaqun
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: of Protease:Inhibitor Complexes
FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR FILING DATE: 1999-01-08
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                            240
                                                                                                                                                                                                                                                                                                                    100
                                                                                                                                                                                                                                                                                                                                              300
                                                                                                                                                                                                                                                                                                                                                                      120
                                                                                                                                                                                                                                                                                                                                                                                              360
                                                                                                                                                                                                                                                                                                                                                                                                                        140
                                                                                                                                                                                                                                                                                                                                                                                                                                               420
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 480
                                                                                                                                                                                        120
                                                                                                                                                                 0.4
                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                    80
                                                                                                                                        9
                                                                                                                                                               141 LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                       US-09-965-594-22 (1-197) x US-09-965-594-21 (1-594)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 19, Application US/09965594; Patent No. US20020106642Al; GENERAL INFORMATION:
2.28e-108
1010.00
99.49%
99.49%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 19
LENGTH: 594
```

```
06e-105
                                 virus
                                                                              980.00
97.46%
96.45%
96.46%
                                 Ü
                     ; TYPE: DNA
; ORGANISM: Hepatitis
US-09-965-594-17
                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 5
US-09-965-594-15
                                                            Alignment Scores:
   SEQ ID NO 17
LENGTH: 594
                                                                                                                                                                          7
                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                161
                                                                                                                                                         ٦
                                                                    Pred. No.:
                                                                               Score:
                                                                                                                                                         ò
                                                                                                                                                                           셤
                                                                                                                                                                                                             q
                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                          Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 17, Application US/09965594

Patent No. US20020106642a1

GENERAL INFORMATION:
APPLICANT: Wittekind, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Coldiarb, Valentina
APPLICANT: Coldiarb, Valentina
APPLICANT: Goldfarb, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: Of Protease:Inhibitor Complexes
FILE REFERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594

CURRENT FILING DATE: 2001-09-27

PRIOR APPLICANTION NUMBER: 60/115,271

PRIOR PRIOR FILING DATE: 1999-01-08

NUMBER OF SEQ ID NOS: 26

SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                  480
                                                                                                                                                                                                                                                                                                                                                                                                                                             61 TACGCTCAGCAGACTCGAGGTGAGGAGGTTGCCAAGAAACCTCCCAGACCGGTGGTGAC 120
                                                                                                                                                                                                                                  180
                                                                                                                                                                                                                                                                                       GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr 120
                                                                                                                                                                                                                                                                                                                                         LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys 160
                                                                                                                                                                                                                9
                                                                                                                                      50
                                                                                                                                                        9
                                                                                                                                                                          40
                                                                                                                                                                                                               MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                               TyralaglnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                            CTGCTGTCCCCGCGTCCCGATCTCCTACCTGAAAGGTTCCTCCGGTGGTCCCCTGCTGTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 594
2 2 2 0 0 0
                                                 Length:
Matches:
Conservative:
Mismatches:
                                                                                                                   US-09-965-594-22 (1-197) x US-09-965-594-19 (1-594)
                                                                                       Indels:
Gaps:
                                               1.27e-106
995.00
98.98%
97.97%
           ; ORGANISM: Hepatitis C virus US-09-965-594-19
                                                                              Best Local Similarity:
Query Match:
                                                                    Percent Similarity:
                                      Alignment Scores:
Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09-965-594-17
                                                                                                                                                                          21
                                                                                                                                                                                                                                                                                                                                                                                     361
                                                                                                                                                                                                                                                   61
                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                                                              101
                                                                                                                                                                                                                                                                                                                                                                   121
                                                                                                                                                                                                                                                                                                                                                                                                         141
                                                                                                                                                                                                                                                                                                                                                                                                                          421
; TYPE:
                                                                                                                                                     g
                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                    à
                                                                                                                                                                         õ
                                                                                                                                                                                       đ
                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                   οχ
                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                             δy
```

```
Sequence 15, Application US/09965594
Patent No. US20020106642A1
GENERAL INFORMATION:
APPLICANT: Withekind, Michael
APPLICANT: Weinhelmer, Steven
APPLICANT: Weinhelmer, Steven
APPLICANT: Zhang, Yaqun
APPLICANT: Zhang, Yaqun
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies
TITLE OF INVENTION: of Protease:Inhibitor Complexes
TITLE OF INVENTION: Of Protease:Inhibitor Complexes
TITLE PEPERENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
                                                                                                                                                                                      ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                         80
                                                                                                                                                            9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                       181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
594
22
55
0
                          Conservative:
Mismatches:
Indels:
                                                                                                US-09-965-594-22 (1-197) x US-09-965-594-17 (1-594)
Length:
Matches:
```

```
Alignment Scores:
                                                                                                                                                                                                   21
                                                                                                                                                                                                                                                                                                                                                       301
                                                                                                                                                                                                                                                                                                                                                                        121
                                                                                                                                                                                                                                                                                                                                                                                        361
                                                                                           ..
No.:
                                                                                                    Score:
                                                                                           Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                   음
                                                                                                                                                                                                                                                                                                                                    ç
                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                    ò
                                                                                                                                                                                  g
                                                                                                                                                                                                   ò
                                                                                                                                                                                                                   đ
                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                  셤
                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 25, Application US/09965594
Patent No. US20020106642A1
GENERAL INFORMATION:
GENERAL INFORMATION:
APPLICANT: Weinhelmer, Steven
APPLICANT: Weinhelmer, Steven
APPLICANT: Colidato, Valentina
APPLICANT: Colidato, Valentina
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Acailitating Inhibitor Screening and Structural Studies
TITLE FEFRENCE: DB17Sequences
GILE REFRENCE: DB17Sequences
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT-FILING DATE: 2001-09-27
                                                                                                                                                                                                                                                                                                                                                                                         160
                                                                                                                                                                                                                                                                                                                                                                                                                        100
                                                                                                                                                                                                                                                                                               LeuValThrargHisalaaspvalileProvalargargargGlyaspSerargGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                 180
                                                                                                                                                                                                                                                                      240
                                                                                                                                                                                                                                                                                                                                                                       361 crestraccestracersacerrarecessitererestesisasarecessiste 420
                                                                                                                                                                                                                                                     80
                                                                                                                                                                                                                    9
                                                                                                                                                    20
                                                                                                                                                                   9
                                                                                                                                                                                  40
                                                                                                                                                                                                                                                    TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                      SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp
                                                                                                                                                                                                                                                                                                                                                                                         LeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                         594
187
2
8
0
0
                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                    US-09-965-594-22 (1-197) x US-09-965-594-15 (1-594)
                                                                          6.71e-103
963.00
95.94%
94.92%
94.78%
                                           C virus
HUMBER OF SEQ ID NOS: 26

SOFTWARE: Patentin Ver. 2

SEQ ID NO 15

LENGTH: 594

TYPE: DNA

TYPE: DNA

US-09-965-594-15
 26
                                                                                         Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 6
US-09-965-594-25
                                                                                                                                                                                    21
                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                      181
                                                                                                                                                                                                                                                                                       81
                                                                                                                                                                                                                                                                                                                                                        121
                                                                            No.:
                                                                                     Score:
                                                                                                                                                                                                                                                                                                                                                                                                                         0 y 0 y
                                                                                                                                                                                                                                                                                                                                                                                          60 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                     ò
                                                                                                                                                                 q
                                                                                                                                                                                  ò
                                                                                                                                                                                                  Oy
Oy
                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                 o
O
                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                  á
                                                                                                                                                                                                                                                                                                                                   엄
                                                                                                                                                                                                                                                                                                                                                       δ
```

```
Sequence 13, Application US/09965594
Patent No. US20020106642A1
GENERAL INFORMATION:
APPLICANT: Withekind, Michael
APPLICANT: Weinheimer, Steven
APPLICANT: Zhang, Yaqun
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Pacilitating Inhibitor Screening and Structural Studies
FILE REFERENCE: DB17Sequences
FILE REFERENCE: DB17Sequences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      420
                                                                                                                                                                                                                                                                                                                                                                                                          100
                                                                                                                                                                                                                                                                                                                                                                                                                                300
                                                                                                                                                                                                                                                                                                                                                                                                                                                     120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            360
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              160
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    480
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               540
                                                                                                                                                                                                                                                                                                                                                                                    240
                                                                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                             120
                                                                                                                                                                                                                                                                                                                  9
                                                                                                                                                                                                                                                                                                                                                               80
                                                                                                                                                                                                                            20
                                                                                                                                                                                                                                                 9
                                                                                                                                                                                                                                                                      0
                                                                                                                                                                                                                                                                                                                                                            101 GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CGGCTCCGCAGGGTTCCCGTGACCCGTGCACCTGCGGTTCCTCCGACCTGTAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    594
185
0
0
0
                                                                                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                        US-09-965-594-22 (1-197) x US-09-965-594-25 (1-594)
                                                                                                                           Length:
Matches:
                                                                                                                                                                                   Gaps:
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PATENTIN VET. 2.0
SEQ ID NO 25
LENGTH: 594
TYPE: DNA
                                                                                                                          9.3e-100
936.00
93.91%
93.91%
92.13%
                                                                                C virus
                                                                             ; ORGANISM: Hepatitis
US-09-965-594-25
                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-13
```

```
Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-10-133-133A-6
                                                                                         LENGTH: 588
TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                 61
                                                                                                                                                                                                                                                                                                                                                                                                                             235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          475
                                                                                                                                                                                                                                                                                                                           41
                                                                                                                                                                                                                                                                                                                                                                                    175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               181
                                                                      SOFTWARE: E
                                                                                                                                                                                              Query Match
                                                                                                                                                      ..
Q
                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: Wittekind, Michael
APPLICANT: Weinhelmer, Steven
APPLICANT: Zhang, Yaqun
APPLICANT: Chang, Yaqun
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Modified Forms of Hepatitis C NS3 Protease for
TITLE OF INVENTION: Facilitating Inhibitor Screening and Structural Studies
                                                                                                                                                                                                                                                                                                                                                                                                                           LeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArgGlySer 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               161 ProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       534
                                                                                                                                                                                                                                                                            234
                                                                                                                                                                                                                                                                                                                                              80
                                                                                                                                                                                                                           20
                                                                                                                                                                                                                                              54
                                                                                                                                                                                                                                                                  40
                                                                                                                                                                                                                                                                                                       41 LysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla 60
                                                                                                                                                                                                                                                                                                                                             MetLysLysLysGlySerValValIleValGlyArgIleAsnLeuSerGlyAspThrAla
                                                                                                                                                                                                                                              1 ATGAAAAAAAAAAGGATCCGTTGTTATCGTCGCCCTATAGTACTGAACGGT-----GCT
                                                                                                                                                                                                                                                                  TyralaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerH1sThrGlyArgAsp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           181 AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    588
183
9
1
                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                        US-09-965-594-22 (1-197) x US-09-965-594-13 (1-588)
                                                                                                                                                                          Indels:
                                                                                                                                                                                    Gaps:
CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 4, Application US/09965594 Patent No. US20020106642Al GENERAL INFORMATION:
                                                                                                                                5.98e-99
929.00
94.42%
92.89%
                                                                                 ; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-965-594-13
                                                                                                                                         Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
US-09-965-594-4
                                                                       LENGTH: 588
                                                                                                                                                                                                                                                                  21
                                                                                                                                                                                                                                                                                      55
                                                                                                                                                                                                                                                                                                                                                                 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              415
                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                   Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                            ò
```

```
161 ProalaGlyHisalaValGlyIlePheargalaAlaValSerThrArgGlyValAlaLys 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          81 SerProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             55 TACGCTCAGCAGACTCGAGGTCTGCTGGGTTGCATCACCACCTCCCTGACCGGTGGTGAC 114
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ThrSerIleAsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CCGGCTGGTCACGCTGTTGGTATCTTCCGTGCTGTTTGCACCCGTGGTGTTGCTAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LysasnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AlaValAspPheIleProValGluSerLeuGluThrThrMetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: CATHERS, Brian
APPLICANT: NEUTEBOON, SASKIA
APPLICANT: SHEPARD, Michael
TILLE OF INVENTION: VIRAL ENZYME ACTIVATED PROTOTOXOPHORES
                                                                                                                                                                                                                                                                         588
181
1
13
2
       Complexes
                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-22 (1-197) x US-09-965-594-4 (1-588)
                                                                                                                                                                                                                                                                         Length:
Matches:
TITLE OF INVENTION: of Protease:Inhibitor of FILE REFERENCE: DB17Sequences CURRENT APPLICATION NUMBER: US/09/965,594
CURRENT FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/115,271
PRIOR FILING DATE: 1999-01-08
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                  Indels:
     of Protease: Inhibitor
                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; Sequence 6, Application US/10133133A
; Publication No. US20030114385A1
; GENERAL INFORMATION:
                                                                                                                                                                                                    ; ORGANISM: Hepatitis C virus US-09-965-594-4
                                                                                                                                                                                                                                                                                        902.00
92.39%
91.88%
88.78%
                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
```

9

20

54

80

140

160

```
RESULT 11
US-09-238-076-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyProLeuLeuCysPro 161
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     498
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      558
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              162 AlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyValAlaLySAla 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GCCCAGCAGACAAGGGGCCTCCTAGGGTGCATAATCACCAGCCTAACTGGCCGGGACAAA 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AlaGInGInThrArgGlyGluGlnGlyThrGlnLySThrSerHisThrGlyArgAspLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SerIleAsnGlyValLeuTrpThrValTyrH1sGlyAlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThr
TITLE OF INVENTION: AND USE OF SAME TO TREAT VIRAL INFECTIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     182 ValAspPheIleProValGluSerLeuGluThrThrMetArqSer 196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  612
174
3
15
15
                                                                                                                                                                                                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-22 (1-197) x US-10-133-133A-6 (1-612)
          FILE REFERENCE: NB 2021.00
CURRENT APPLICATION NUMBER: US/10/133,133A
CURRENT FILING DATE: 2002.04-26
PRIOR APPLICATION NUMBER: 60/286,983
PRIOR FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 21
SEQ ID NO 6
SEQ ID NO 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 3, Application US/09742659; Patent No. US20010034019A1
SEREAL INFORMATION:
APPLICANT: HONG, Zhi
APPLICANT: Zhong, Weldong
APPLICANT: Zhong, Weldong
APPLICANT: LIGHTAN APPLICANT: 
                                                                                                                                                                                                                                                                                                                                                                      2.02e-91
864.50
90.77%
89.23%
85.09%
                                                                                                                                                                                                                                                                           ORGANISM: Hepatitus C. Virus
                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                           612
                                                                                                                                                                                                                                                                                                     US-10-133-133A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
US-09-742-659-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   142
                                                                                                                                                                                                                                                   TYPE: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    42
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  259
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    379
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                           LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DP
DP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ογ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ω
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
3593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3773
                                                                                                                                                                                                                                                                                                                                                          3473
                                                                                                                                                                                                                                                                                                                                                                           3653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                       73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 93
                                                                                                                                                                                                                                                                                                                                    33
                                                                                                                                                                                                                                                                                                                                                 3354 CGTAGGGGCCAGGAGATACTGCTTGGACCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                       54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ArgGlyAspSerArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySer
                                                                                                                                                                                           9646
168
10
17
                                                                                                                                                                                             Length:
Matches:
Conservative:
                  APPLICANT: Lean, Stanley M.
TITLE OF INVENTION: Chimeric HCV/GBV-B viruses; FILE REFERENCE: IDO1116
FURENT APPLICATION NUMBER: US/09/742,659
CURRENT APPLICATION NUMBER: US 60/171,469
PRIOR APPLICATION NUMBER: US 60/171,469
PRIOR FILING DATE: 1999-12-22
NUMBER OF FEL ID NOS: 16
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
                                                                                                                                                                                                                                                                                         3 LysLysGlySerValVallleValGlyArglleAsn-
                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x US-09-742-659-3 (1-9646)
                                                                                                                                                                                                                             Mismatches:
                                                                                                                                                                                                                                       Indels:
Gaps:
Wright-Minogue, Jacquelyn
Lau, Johnson Y.
                                                                                                                                                                                            5.25e-88
848.50
87.25%
82.35%
                                                                                                                                         ; TYPE: DNA
; ORGANISM: Hepatitis C virus
US-09-742-659-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1111111111
3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MetArgSerPro 197
                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                   Alignment Scores:
Pred. No.:
                                                                                                                               LENGTH: 9646
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        154
                                                                                                                                                                                                                                                                                                                                    15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             194
 APPLICANT:
                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                            ó
                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             줨
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     줨
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
```

```
3894 IGCACCCGIGGAGTGGCTAAGGCGGTGGACTITAICCCTGTGGAGAACCTAGAGACAACC 3953
                                                                                                                                                                                                                                                        94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                         3774 CGAGGTGATAGCAGGGGTAGCCTGCTTTCGCCCCGGCCCATTTCCTACTTGAAAGGCTCC
                                                                                                                                                                                                                                                                                                             174 SerfhrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr
                                                                                            114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                       154 SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (RCV) AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 0S/09/995,937
FILING DATE: 28-No. US20030028010A1-2001
CLASSIFICATION: <UNKNOWN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,756
FILING DATE: 04-May-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1, Application US/0995937; Publication No. US20030028010A1; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEPHONE: 314-727-5188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                        3954 ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY: 11near MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                                                                                                                   194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY: ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANT I - SENSE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-995-937-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-995-937-1
                                                                                                                                                                                                                                                                     đ
                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                   용
                                                                                          ò
                                                                                                                           a
                                                                                                                                                                 5
                                                                                                                                                                                                   a
                                                                                                                                                                                                                                       ó
                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                   à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3354 CGTAGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys 33
                                       APPLICANT: RICE, CHARLES et al.

TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9646
168
10
17
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3 LysLysGlySerValVallleValGlyArglleAsn......
                                                                                                                                                                                                                                                                       MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERAITMG SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-22 (1-197) x US-09-238-076-1 (1-9646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels:
                                                                                                                                        ADDRESSEE: HOWELL F HAFERKAMP, L.C.
STREET: 7733 FORSYTH BLVD., SUITE 1400
CITY: ST. LOUIS
                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
                                                                                                                                                                                                                                                                                                                                                                                                                                                    FTLING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727.5188
Sequence 1, Application US/09238076
Patent No. US20020102540A1
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5.25e-88
848.50
87.25%
82.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LENGTH: 9646 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
                                                                                                                                                                                                                                                      COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                      CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity:
                                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOLECULE TYPE:
HYPOTHETICAL: N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similari
                                                                                                                                                                                                                                   63105
                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                       ç
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-238-076-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
```

```
3773
                                                                                                                                                                                                                                                                                                                                                                                                                                                            3533
                                                                                                                                                                                                                                                                                     3593
                                                                                                                                                                                                                                                                                                                                                                          3834 TCGGGGGGTCCGCTGTTGTGCCCCGCGGGACACGCCGTGGGCCTATTCAGGGCCGCGGGTG 3893
                                                                                                                                      3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                              113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             173
                                                                                                                                                                                                                                                                                                                                                                                                              114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            174 SerThrArgGlyValAlaLySAlaValAspPhcIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                               53
                                                                                                                                                                                                                                                              73
                                                                                                                                                              33
                                                                                                                                                                                                                            AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SerGlyGlyFroLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAalaVal
                                                                                                                                                              ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys
                                                                                                                                                                                                               ThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                              54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                      TGCGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCY) AND USES THEREOF
9646
168
10
17
17
                           Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSYTH BIVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
                                                                                       JS-09-965-594-22 (1-197) x US-09-995-937-1 (1-9646)
                                                                                                                LysLysGlySerValValIleValGlyArgIleAsn-
  Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
WEDDUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           et al
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 1, Application US/09917563; Publication No. US20030073080A1; GENERAL INFORMATION: APPLICANT: RICE, CHARLES et
 5.25e-88
848.50
87.25%
82.35%
83.51%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MetArgSerPro 197
                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 13
US-09-917-563-1
                                                                                                                                                                                                                                                                                                                                    3594
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3894
                                                                                                                ~
                                                                                                                                                                                                               34
                                                                                                                                                                                                                                                                                     3534
                                                                                                                                                                                                                                                                                                              74
                                                                                                                                                                                                                                                                                                                                                                                     3654
                                                                                                                                                                                                                                                                                                                                                                                                                                      3714
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           194
                                                                                                                                                                                                                                                                                                                                                             94
  Pred. No.:
                Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                   Q
                                                                                                                                                             ò
                                                                                                                                                                                 g
                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                   셤
                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                  đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                à
```

```
3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3533
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3593
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3653
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3713
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3773
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3833
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        153
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            114 CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCGGAATGGTCTCCAAGGGGTGGAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrargGlyGluGlnGlyThrGlnLys
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3414 TTGCTGGCGCCCATCACGGCGTACGCCCAGCAGAGGAGAGGCCTCCTAGGGTGTATAATC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 54 AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              34 ThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3654 GACCAAGACCTTGTGGGCTGGCCCGGTCCTCAAGGTTCCCGGTCATTGACACCCTGCACC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3714 TGGGGCTCCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGCGCGG
                                                                                                                                                                                                                                                                                                                                                                9646
168
10
17
9
                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 LysLysGlySerValVallleValGlyArgIleAsn---
                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-22 (1-197) x US-09-917-563-1 (1-9646)
           APPLICATION NUMBER: US/09/917,563
FILING DATE: 27-Jul-2001
CLASSIFICATION: <urknown>
                                                                                                                                                                                                                                                                                                                                                                                                                       indels:
                                                   PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/238,076
FILING DATE: 26-JAN-1999
                                                                                                                                                                                                                                                                                              ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
REGISTRATION NOWBER: 35,197
REFERENCE/DOCKET NUMBER: 605
TELECOMMUNICATION INFORMATION:
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                      TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9646 base pair
TYPE: nucleic acid
STRANDEDNESS: double
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                5.25e-88
848.50
87.25%
82.35%
83.51%
                                                                                                                                                                                                                                                        TOPOLOGY: linear MOLECULE TYPE: CDNA
                                                                                                                                                                                                                                                                                   HYPOTHETICAL: NO
                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                        US-09-917-563-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               154
                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                  ..
Q
                                                                                                                                                                                                                                                                                                                                                                               Score:
                                                                                                                                                                                                                                                                                                                                                              Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                줨
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       셤
```

33

ò g ò g

```
3834 resesses receserations consequences and a resesses recesses r
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   3534 GCTACCCAAACCTTCCTGGCAACGTGCATCAATGGGGTATGCTGGACTGTCTACCACGGG 3593
                                                                                                                                                                                                                                                                                                                                                             174 SerThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                            93
                                                                                                                                                                                                                                                                                                                           94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SerGlyGlyDroLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal
                                                                                      AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly
                                                                                                                                                                                                        74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IIILE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C VIRUS (HCV) AND USES THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/995,937
FILING DATE: 28-No. US20030028010a1-2001
CLASSIFICATION CORNOWD>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/034,756
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STREET: 7733 FORSYTH BLVD., SUITE 1400 CITY: ST. LOUIS
STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REGISTRATION NUMBER: 35,197
REFERENCE/DOCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 5, Application US/09995937; Publication No. US20030028010A1
GENEAL INFORMATION: GENERAL INFORMATION: APPLICANT: RICE, CHARLES et al.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 04-May-1998
ATTORNEY/AGENT INFORMATION:
NAME: HOLLAND, DONALD R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NCE CHARACTERISTICS:
LENGTH: 12980 base pairs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEPHONE: 314-727-5188
TELEFAX: 314-727-5092
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
ADDRESSEE: HOWELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         194 MetArgSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        COUNTRY: US
ZIP: 63105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
US-09-995-937-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  154
                                                                                      54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                           qq
                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                         ď
                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                         9
                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3414 TTGCTGGCGCCCATCACGGCGTACGCCCAGCAGACGAGGCCTCCTAGGGTGTATAATC 3473
                           3894 TGCACCCGTGGAGTGGCTAAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGAACA 3953
174 SerThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThr 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ---LeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLys 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34 ThrSerHisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr 53
                                                                                                                                                                                                                                                                                                                                                                                  APPLICANT: RICE, CHARLES et al.
TITLE OF INVENTION: FUNCTIONAL DNA CLONE FOR HEPATITIS C
TITLE OF INVENTION: VIRUS (HCV) AND USES THEREOF
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/238,076
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LysLysGlySerValValIleValGlyArglleAsn---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-22 (1-197) x US-09-238-076-5 (1-12980)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: HOWELL & HAFERKAMP, L.C.
STREET: 7733 FORSTTH BLVD., SUITE 1400
CITY: ST. LOUIS
STATE: MO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: HOLLAND, DONALD R.
REGISTRATION NUMBER: 35,197
REFERENCE/POCKET NUMBER: 6029-4831
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEPRAX: 314-727-6092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/034,756
FILING DATE:
                                                                                                                                                                                                                                                                                            Sequence 5, Application US/09238076 Patent No. US20020102540Al GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TELEFAX: 314-727-6092
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.72e-88
848.50
87.25
82.35
83.51
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      : 12980 base pairs
nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ATTORNEY/AGENT INFORMATION
                                                                                                                                                    3954 ATGAGATCCCCG 3965
                                                                                                                     MetArqSerPro 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   double
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ZIP: 63105
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGI.

MOLECULE TYPE: CD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ANTI-SENSE: NO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COUNTRY:
                                                                                                                                                                                                                                       RESULT 14
US-09-238-076-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-238-076-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15
                                                                                                                     194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
```

ö 셤 ŏ

```
3354 CGTAGGGGCCAGGAGATACTGCTTGGGCCAGCCGACGGAATGGTCTCCAAGGGGTGGAGG 3413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           74 AlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsnVal 93
                                                                                                                                                                                                                                                                                                                                                                                                                                                    7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AlaThrGlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly 73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LysLysGlySerValValIleValGlyArgIleAsn------------
                                                                                                                                                                                                                                    12980
168
10
17
9
                                                                                                                                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-22 (1-197) x US-09-995-937-5 (1-12980)
                                                                                                                                          5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Search completed: August 31, 2003, 04:55:02 Job time : 189.482 secs
                                                                                                                                             ..
02
                                                                                                                                             Ω
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
SEQUENCE DESCRIPTION: SEQ I)
                                                                                                                                                                                                                                      7.72e-88
848.50
87.25%
82.35%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ATGAGATCCCCG 3965
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MetArgSerPro 197
                                                                                                                                                                                                                                                               Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                Alignment Scores:
Pred. No.:
       TYE

STH

TOG

MOLECULE

HYPOTHEN

ANTI-SEN

SEQUENCE

13 SEQUENCE

16 SEQUENCE

17 SEQUENCE

18 SEQUENCE

18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3954
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          154
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         á
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       \delta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    oy
o
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        oy
B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
```

,*

GenCore version 5.1.6 (c) 1993 - 2003 Compugen Ltd. Copyright

	copy. Lync (c) 1530 compages act.	
OM protein - nu	OM protein - nucleic search, using frame_plus_p2n model	
Run ou:	August 30, 2003, 19:20:43 ; Search time 1910.31 Seconds (without alignments) 2506.388 Million cell updates/scc	
Title: Perfect score: Sequence:	US-09-965-594-22 1016 1 MKKKGSVVIVGRINLSGDTAVAKAVDFIPVESLETIMRSP 197	
Scoring table:	BLOSUM62 Xgapop 10.0 , Xgapext 0.5 Ygapop 10.0 , Ygapext 0.5 Fgapop 6.0 , Fgapext 7.0 Delop 6.0 , Delext 7.0	
Searched:	22781392 seqs, 12152238056 residues	

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

45562784

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

em_gss_pro: *
em_gss_rod: *
em_gss_rod: *
em_gss_phg: *
em_gss_vrl: *
gb_gssl: * em_gss_vrt:*
em_gss_fun:*
em_gss_mam:* em_gss_hum:* em_gss_inv:* em_esthum: *
em_estin: *
em_estin: *
em_estin: *
em_estin: *
em_estin: *
em_htc: *
gb_esti: *
gb_est: *
gb_htc: * em_estfun:* em_estom:* em_gss_pln: em_qss_mus: gb_est4:* gb_est5:* em_estba:* gb_est3:* EST:* 255... 1011... Database :

29: qb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	escriptio	OK 66609	787	: =	00000	0000	1972 Manyoeu	15803 AGENCOU	53244 963042C	06704 PUHKD12	06705 PUHKD12	04699 6018882	42381 ic83bl	10673 T3 end	57279 AGENCOU	12566 SLA005F	8021 RPCI-11	1177	465	728	583	605	024	980	877	33.7	833	B1459445 603200433	100	4000	116	911	232	926	9969	1043	692	575	345	662	054	18007 UI-E-CI	5317 K-EST009	2750 60129841	105	0847 K-FC#010	OCAL P-ESTOTO
SOMERIES	۵	: 02	2	B0926101	ď	2	, 4	2:	2 ,	٤:	٩.	₹.	⊆.	9.	~	~	8	コ	7	7	ŭ	9	2	2	9	ກຸ	9	B1409445	0	7		5	2	õ	8	2	9	Ñ	33	9	2	708	25	382	M961	330	2
		14		13	000	1 -	? .	7;	26	57	2	10	53	5	N	N	œ	m	m	σ	0		80	0	0	٧.	4 (7 .	٠,	٠,	2	~	σ	6	_	m	m	m	~	m	Н						
	Length	1031	1199	1403	644	000	1	1140	107	7/7	687	984	701	1062	1440	528	260	1213	938	1026	617	886	528	958	1035	1/33	960	753	2 4	908	917	586	670	812	814	817	824	878	920	1106	1141	200	569	290	631	658) }
ď	Query Match	'n	6	5	٧		•	•	•	•	•	₹.	*	m.	m.	?	?	۲.	۲.	٦.	4	7	0	0.	9	٠,	? <		? <	9	0	σ.	σ.	σ.	σ.	o.	σ.	ō.	σ.	σ.	6	σ.	o.	6	σ.	σ	•
	Scor	104.	10	100.5	σ	2, 70	٠.	^ נ	n .	U. U.	٠ ،	ς,	σ.	٠	•	66	93	σ	92.5	ς.	92	σ,	Ξ,	∹.	21.5	∹ '	7.6	16	10	10	16	0	。	90.5	٥.	。	。	。	。	0	。	06	90	90	06	06	,
	sult No.		8			٠,٠				စ (ပ	٠	د 10	٦,	12	13	4	٦.	د 1 ₆	17	18	13	50	~	~ (573	4 (67	27		i	30	m	٣	c 33	m	m	m	m	m	(F)	40	4	4	4	c 44	4	•

EST 29-APR-2003 EST.

Mus musculus (house mouse)

Mus musculus

Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutharia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1031) CB950999 1031 bp mRNA linear EST AGENCOURT_13445496 NIH_MGC_177 Mus musculus cDNA clone IMAGE:30316162 5', mRNA sequence. CB950999.1 GI:30205777 DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM RESULT 1 CB950999

REFERENCE

ALIGNMENTS

~

```
175 rArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMetAr 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1199
50
24
80
59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  211 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1. .1199
/organism="Homo saplens"
/nol_type="mRNA"
/db_xref="taxon:9606"
/clone="!MAGE:6192708"
/sex="male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-22 (1-197) x BQB92487 (1-1199)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  343 g
                                                                                                                                                                                                               BQ892487.1 GI:22284501
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Technologies.
362 c
                                                                                                                                                                                                                                                  Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              101.00
34.74%
23.47%
9.94%
                                                                                195 gSerPro 197
                                 772 C-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                                                                                                                                                                                                                                AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                  REFERENCE
                                                                                                                          RESULT 2
BQ892487
                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                              õ
                                                                                                                                                                                                                                                                                                          543
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             104 Gln-------GlySerArgSerLeuThrProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           119 LeuTyrLeuValThrArgHisAlaAsp~ValIleProValArgArgArgGlyAspSerAr 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             651
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          138 gGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGly----- 155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -GlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerTh 175
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        395 ATTCAAGGCTATCCCCAAAACAAGAGAGTACATCGGCAGCTTTTCCTT---CACTCATTT 451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   84 GlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaPro 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          652 AAACNCCCTTAACCGTCCCTCCCAACAATATTTCTTCAAGCGTCCTCTGATTTCCCTAA 711
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           44 ValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIle 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      64 AsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys 83
                                                         Email: cgapbs:remail:nh.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Michael Brownstein
cDNA Library Arrayed by: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
Plate: NDCMIOT row: b column: 11
High quality sequence stop: 333.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               604 ------ACCACCCATGGGGCTTGTTGTTCCCGGCTCTCCCGTGGGCAATTACA
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1031
51
16
62
35
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                   1. .1031
/organism-"Mus musculus"
                                               Contact: Robert Strausberg, Ph.D.
                                                                                                                                                                                                                                                               /mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30316162"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-22 (1-197) x CB950999 (1-1031)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3.53
104.50
41.10%
31.29%
10.29%
                                  Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No.:
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                      FEATURES
```

qq ò g ò q ò g ò 셤 õ

셤 õ

```
//Clone_lib="Lupski_sympathetic_trunk"
//clone_lib="Lupski_sympathetic_trunk"
//note="Vector: pcMv-SPORT6 (Life Technologies); Site_l:
Not1; Site_2: Sali; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCGTCCG-3' and
5'-TCGACCCAGCGTCCG-3' and
5'-GACTAGTTCTAGATCGCGAGCGCCCCT(15)-3'. Size selected >
1' bx for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
                                                                                                                                                                                                                                                                 1199 bp mRNA linear EST 16-AUG-2002 mRNA linear EST 16-AUG-2002 IMAGE:6192708 5', mRNA sequence.
GCAGGAGAGAACCTTACCCCAAACAG------AAGGCA 369
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16 SerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSer 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1199)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Techhologies, Inc.
CDNA Library Preparation: Life Techhologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC.clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LiAMissys row: column: 13
High quality sequence start: 57
High quality sequence stop: 394.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28 others
```

Page

쥠

g ò

ö

g ò 셤 ò g ò

ò

```
1207 GCGGCGGCGTGGCCGAGTATATACCGCCGAGTGCGAACCCCACGGGCGCGTGGAACGTTG 1148
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1381 AGAGGGTGTGTTCANCGGTCAGGACAGGTC---GCCGCACTCGACGCGTCGGCCAGAG 1325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1252 CAGGGGGTA---TGGTTATCACGGGCTGGGGCAGGTACT------TCCCCTAAA 1208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSS 29-JAN-2003
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1324 ACTTGTCGGGGGGCGCGTTGGCGCATACCCCGGGTCGGATCGAGGTCAGCGCCGCTTGTAT 1265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      264 ACAGAGGGAAA----- 1253
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    84 --GlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaP 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              103 roGlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValT 123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                123 hrargHisAlaAspValIleProValArgArgGlyAsp------- 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1030 CGCCCGCCGCGCGCGAAC------AGGGTAATAATCATATGCGGCGGGGAT 983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             154 SerGly-----GlyProLeuLeu 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        982 TTCGCATTCCGGCGGGAGAAGGCGGCGGTCGGGGGGCGCGCGGGCCTGCGGGCCGCTGCAGG 923
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSS.
Danio rerio (zebrafish)
Danio rerio
Danio rerio
Bukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygil; Neopterygil; Teleostei; Ostariophysi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  160 CysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGlyVal 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /clone="INAGE:6335718"
/lab_host="DH10B (phage-resistant)"
/clone=llb="NH4MGC_13"
/clone=llb="NH4MGC_13"
/note="Organ: otocysts: Vector: pcMv-SPORT6.1.ccdb;
Site_1: EcoRV; Site_2: Not1; Cloned unidirectionally.
Primer: oligo dT. Average insert size 1.95 kb.
Constructed by ResGen, Invitrogen Corp. Note: this is NH4MGC Library."

521 c 237 g 345 t 3 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ых 23 у 1 рим 1 near GSS 29-JAN-Danio rerio genomic clone DKEY-283L13, genomic survey sequence. 
В X238988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4 LysGlyServalValIleValGlyArgIleAsnLeuSerGlyAspThrAlaTyrAlaGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24 GlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThrGlyArgAspLysAsnGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1147 ACCAA---GAGAGGCACCTGACGCCCCTCCCCTGTGGGCNTGTCGATAATAACAAATGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        137 -----SerArgGlySerLeuLeuSerProArgProlle-SerTyrLeuLysGlySer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            44 ValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIle
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64 AsnGlyValLeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLys
                                                                                                                                                                                                                                                                                                                          Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-22 (1-197) x BQ926101 (1-1403)
/db_xref-"taxon:10090"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BX238988.1 GI:28161322
                                                                                                                                                                                                                                                                                                      100.50
35.32%
27.86%
                                                                                                                                                                                                                                                                                                                                                                      9.838
                                                                                                                                                                                                                                                                                                                        Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                      Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 4
BX238988/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
                                                                                                                                                                                              BASE COUNT
                                                                                                                                                                                                                                                                                        ..
9
                                                                                                                                                                                                                 ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AGENCOURT_8752655 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6335718 B0926101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    969
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  142
                                                                                                                                            90 TyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeu 109
                                                                                                                                                                                                                                                                                                                                                                    636
                                                                                                                                                                                                                                                                                                                                                                                                                                                         121 -LeuValThr-------ArgHisAlaAspValIleProValArg------ 132
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         143 -SerProArgProlleSerTyrLeuLys-------GlySerSerGlyGlyPr 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               757 CCACCCTCGGCCCACCGATGCGCTAAGCCTCCCTTTACAAGCCACCGCCGCGGCCCCCC 816
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            oLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGl 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    817 CCTAACATCTCCTACCCCTGCGGCGCGCGGGGGGGGGGCGCGTGGGGCGCATACGGGCTCAGGG 876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1403)
                                                         370 CATGGGGGAAATCGCCCGCTTCAGAGAGGAGGTCGTTCATGTTTTCTGAAAACATAACCG 429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman,
Ph.D.
                                                                                                     54 AlaThrGlnThrPheLeu------AlaThrSerIleAsnGlyValLeuTrpThr
                                                                                                                                                                                                                                                                                                                                                                                                             577 GCTGCCTGCACACCACCCCATGACCGATGTTGGGCCTTATGTGGAAACGGCGAGGCGTTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ------ArgargGlyAspSerArgGlySerLeuLeu-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     þe
                                                                                                                                                                                          ValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProValThrGlnMet
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NIH-WGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can found through the I.M.A.G.E. Consortium/LLNL at:
           HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThr---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                cDNA Library Preparation: ResGen, Invitrogen Corp
                                                                                                                                                                                                                                                                                                                        550 TTTACCACATGIGACAGIGACTTT-----------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             177 yvalAlaLysAlaValAspPheIleProValGluSer 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           High quality sequence stop: 101. Location/Qualifiers

    1403
    /organism-"Mus musculus"
    /mol_type-"mRNA"

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mus musculus (house mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BQ926101.1 GI:22341132
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mus musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              133
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      877
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sonrce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 3
BQ926101/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURES
```

셤

g ò g οy g

ò

q

ò

```
Tumor Gene Index
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              74
                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                              VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                 AUTHORS
TITLE
                                                                                                                                                                                               JOURNAL
COMMENT
              ACCESSION
                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  BG089727 62-JAN-2001 mRNA linear, EST 26-JAN-2001 mab90e06.xl NCI_CGAP_Sp2 Mus musculus cDNA clone IMAGE:3977578 3' similar to SW:GRAD_MOUSE P11033 GRANZYME D PRECURSOR;, mRNA
            Humphray, S.J., Huckle, E. and Durham, J.L.

Humphray, S.J., Huckle, E. and Durham, J.L.

Direct Submission

Submitted (27-JAN-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CBIO 18A, UK. E-mail enquiries:

humquery@sanger.ac.uk Unpublished

This sequence was generated from the T7 end of BAC 283L13. 283L13 is part of the banlokey BAC Library created by R. Plasterk and N.V.

Keygene. Further details:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         108 rLeuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspVa 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              147 -IleSerTyrLeuLysGlySerSerGlyGlyProLeuLeu---CysProAlaGlyHisAl 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       nValAspLysAspLeuValGlyTrpGlnAla---ProGlnGlySerArg-----Se 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       316 TCTTGGAAAGGAC-----TGGAGATCCAGACCGTTGGGGACAAAGGCGGAGGACTC 266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GAGACTGCATGGACCACCGGCGTGCCGACGCTGGACACCATCCTGGCAGAGTGGGAGCTC 377
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      376 regaggeserrereceserenceastanceastageseresteateresteses 317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GACCATGGCTATTACCATCACTGGAGGGGGGGAGCAGCTCGGAATCCGGACGCGCCATCTC 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      196 CTTGAGCGGTGCAGCGGACCGGAAGCCTTCCCTGGCCCGGGTGCAGAGTGCCATTAGGGCC 437
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      58 PheLeuAlaThrSerIleAsnGlyValLeu-TrpThrValTyrHisGlyAlaGlyThrAr 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18 AspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThrGlnLysThrSerHisThr 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      77 gThrIleAlaSerProLysGlyProValThrGlnMetTyr------ThrAs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          265 GTCACCACCATGCCCTTTAAAAAGGAGAAATTATTTGTGATTACCATGGGGAGGAGTTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              128 lileProValArgArgGlyAspSerArgGlySerLeuLeuSerProArgPro----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          205 TEGERGRAGGGETTAGGGAGACGGGAGGCCCTCCCCCGGTGTCTCCTCCTACATG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GlyArgAspLysAsnGlnValGluGlyGluValGInIleValSerThrAlaThrGlnThr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 165 aValGlyIlePheArgAlaAlaValSerThrArgGlyVal 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TGCCCCTGTCACCCGGGCATGGAAACCTTCGGAAGGCTTA 49
                                                                                                                                                                                                                                                                                                                                                                                     644
48
25
84
17
                                                                                                                                                                 http://www.sanger.ac.uk/Projects/D_rerio/
Location/Qualifiers
                                                                                                                                                                                         Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps:
Cypriniformes; Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-22 (1-197) x BX238988 (1-644)
                                                                                                                                                                                                                                                                                                                                                                                   8.45
98.00
41.958
27.598
9.658
                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                         Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 5
BG089727/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                           ..
Q
                                                                                                                                                                                                                                                                                                                       BASE COUNT
                                   AUTHORS
                                                                   JOURNAL
                                                                                                                                                                                    FEATURES
                                                                                                                    COMMENT
                                                                                                                                                                                                                                                                                                                                       ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            윱
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
//mol_type="mrna" miscolida
/db_xref="taxon:10090"
/clone="laybac:397578"
/tissue_type="nk cells (flow-sorted)"
/lab_host="DH10B (T1-resistant)"
/clone_lib="NCI_CGAP_SP2"
/note="Organ: spleen; vector: pCNV-SPORT6 (Life
Technologies); mRNA made from flow-sorted NK cells, cDNA
made by oligo-dT priming. Directionally cloned: Average
insert size 1.5 kb. Primary library, non-amplified. cDNA
Library Preparation: David B. Krizman, Ph.D."
                                                                                                                                                                                                                                                                                                                 Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rémail.nih.gov
Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.
Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
http://image.llnl.gov
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          413 ------GGGTCAAGGTCCATCAATGACACTAAAGCATCTGCCCGCCTGCGAGGGTT 363
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAAACATTTCCCATTCCAGATATAAATGCT-----ACTATCTTCTCAGGTGACATC 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                527 ATGCTGTTAAAGCTGGAGAGTAAGGCCAAGAGAACTAAAGCTGTGAGACCCCTCAAGTTG 468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AlaProGlnGlySerArgSerLeu----ThrProCysThrCysGlySerSerAspLeu 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83
                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. I (bases 1 to 629)
NCI-GGAP Http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GlnThrPheLeuAlaThrSerIleAsnGlyValLeuTrpThrValTyrHisGly----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ProLysGlyProValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGln
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----AlaGlyThrArgThrIleAlaSer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1. .629
/organism="Mus musculus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           High quality sequence stop: 422.
High quality sequence stop: 422.
Location/Qualifiers
1. 629
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-22 (1-197) x BG089727 (1-629)
                                                                                Mus musculus (house mouse)
BG089727
BG089727.1 GI:12572290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9.19
97.50
39.23%
26.52%
9.60%
```

: 94- 22 22 98 41 61 02 77	948 AGGGAGGGGCGCGGGGGGGGGGGGGGGGGGGGGCCG	BF863244 BF863244 963042CO2.x1 C. reinhardtii CC-1690, Stress condition I, normalized BF863244. I Cliamydomonas reinhardtii cDNA, mRNA sequence. BF863244. I GI:12253388 BF863244. I GI:1253388 EST. Chlamydomonas reinhardtii Genome: A Model, I (bases I to 70!) Chlamydomonas reinhardtii Genome: A Model, Unicellular System for Analyzing Gene Function and Regulation in Unpublished Contact: Charles Hauser DCMB Box 91000 Duke University Durham, NC 27708-1000 Duke University Durham, NC 27708-1000 Tel: 919 613 8159 Fax: 919 613 8177 Email: chauser@duke.edu.
Query Match DB: US-09-965-5 QY Db 10 QY Db 10 QY QY OY Db 10		RESULT 7 LOCUS LOCUS LOCUS DEFINITION ACCESION VERYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL COMMENT COMMENT FEATURES
120 TyrLeuValThrArgHisAlaAspValIleProValArgArgArgArgArgArgArgArgArgArgArgArgArgA		FEATURES High quality sequence start: 6 High quality sequence start sequence start sequence start sequence start sequence start sequence start sequence se
	RESULT 6 BM915803/C LOCUS DEFINITION ACCESSION VERSION KEYWORS SOURCE ORGANISM AUTHORS TITLE JOURNAL COMMENT	FEATURES source source ORIGIN Alignment Scores Pred. No.: Score: Percent Similari Best Local Similari

us-09-965-594-22.rst

BASE COUNT

ORIGIN

.. Q

ò 셤 ò Q ò q ò QQ õ g ò 셤 ò 8 õ

```
274 GITITCIGCCAG------26AGIAACICGICGCGCA----------245
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu------ 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                393 TGGCCCTACTTGACACGGGTTCAACACATAACTTCATCAACTGCAAGGCGGCA-CAACAG 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       138 ArgGlySerLeuLeuSerProArgProlleSerTyrLeuLysGlySerSerGlyGlyPro 157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       334 CGRGGGGTACTITGGAACCCACACCAGGTCGCCATGTCAAGGTGGCAAATGGAGACCCA 275
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              158 LeuLeuCysproAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGly 177
                                                                                                                                                                                                                                                                                                                                                        /organism="zea mays"
/organism="zea mays"
/mol_type="genomic DNA"
/gtrain="813"
/db_xref="texon:4577"
/clone="zMcBTa469B24"
/note="zMcBTa469B24"
/note="vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"
a 198 c 216 g 208 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    87 ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp------GlnAlaPro 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        507 ACACAGATGATGACCAGG-----AAGAGCAGGACCCAGTGGTTTCCCTACATGCAATAA 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           663 CAGTCTCGTTCAAGCGCCTCACCCCCGCGGAGTGGCGGAACGACGTAACAAGAGGCTAT 604
                      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Bermatophyta; Magnollophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 772)

Whitelaw.C.A., Quackenbush.J., Van Aken.S., Utterback.T., Resnick A., Fraser.C.M. Yuan.Y., San Miguel.P., Ma.J. and Bennetzen.J.
Maize Genomics Consortium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 549 ATCTGGAGG------TCACAGACTTTGCTGACGATGAGATCAAGGGTG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GlnGlyThrGlnLysThrSerHisThrGlyArgAspLysAsnGlnVal ---GluGlyGlu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              48 ValGlnIleValSerThrAlaThrGlnThrPheLeuAlaThrSerIleAsnGlyVal---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            603 GITACAATIGIGACG-----AGCCCIAIGIACGAGGCCACCGCIGICCCAGGCTATICI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 67 LeuTrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----TyrLeuValThrArgHisAlaAspValIleProValArgArgArgGlyAspSer
                                                                                                                                                                                                                    9712 Medical Center Drive, Rockville, MD 20850,
Tel: 301-838-5843
Fax: 301-838-0208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             772
49
24
67
46
10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-22 (1-197) x CC406704 (1-772)
                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                          Email: whitelaw@tigr.org
                                                                                                                                                                                      Contact: Cathy Whitelaw
                                                                                                                                                   Unpublished
Other_GSSs: PUHKD12TD
                                                                                                                                                                                                                                                                                                               ends.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             19
95.50
39.46%
26.49%
9.40%
                                                                                                                                                                                                                                                                                           Seg primer: TR
Class: sheared
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
             Zea mays
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BASE COUNT
           DRGANISM
                                                                                                                                TITLE
JOURNAL
COMMENT
                                                                              REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    οχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            171 CACCACCATACCCTTGCTCTCACGTTGCTCACCAAAATTATGCCCATACGGGCCACTA 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySer 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ArgSerLcuThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAla 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           345 GAGATCATTTGCATGTGGCTTTAGTCACCCCAAGAAGAGCCTGGGAGTGGGCATTTATAA 404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GAAGGGGACGGGAATTCCGTTTGCGGAAAGTGAGCGCCCCCAAGGTCTGACCAAGTGCTA 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyllePhe 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       524
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         170 ArgalaAlaVal-------------------SerThrArgGlyValAlaLysAla--- 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     525 AAAGAAACCATTTGAGTAGGAAGTCGCCGTTTACCCCCGAAGGTGAAGGTCACCTCTAT 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ----GlyProVal 86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ------GlySerLeuLeuSerProArgProlleSer---Tyr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   231 ACAAAGTTACATACACGG------AAGGACCAGCGCGCTTGGCCCACCCCCTTGGAGCCG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AspValIle-----ProValArgArgArgGlyAspSerArg-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CTCCAAGGCAGCAATGGGAGCCTTTCGCGGGTGTGCGGGTGCTCGTCCTCTAATGTCAGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ValAspPheIleProValGluSerLeuGluThrThrMetArg 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GTGGACCGCATACATCTGGAAGACACACGGTGCACTCTACGA 626
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TyrHisGlyAlaGlyThrArgThrIleAlaSerProLys-----
/organism-"Chlamydomonas reinhardtii"
/mol_type="mRNA"
                                                                                                                                                                                                                                                                                                                                                                                                                    701
38
22
63
63
                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-22 (1-197) x BF863244 (1-701)
                                                                                                                                                                                                                                                                                                                                                                                                                  16.8
95.50
38.96%
24.68%
9.40%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                               173
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        285
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               585
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CC406704/c
LOCUS
DEFINITION
```

508

137

ACCESSION VERSION KEYWORDS SOURCE

104 GlnGlySerArgSerLeuThrProCysThrCysGlySerSerAspLeu 11	120 crgggarifagaaggacaccargcagcrgcggggacactgcaggraggacactgc 120TyrLeuValThrargHisAlaAspValIleProValArgArgArgGlyAspSer	DB 380 INSCCEINCINGACAGGGGTTCANCACATCANCACACACACACACACAGGGGGGGGGG	<pre>Qy 158 LeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValSerThrArgGly 177 :: </pre>	<pre>Qy 178 ValalaysalaValAspPhe</pre>	580	RESOLI 10 BF304699/C LOCUS BF304699 984 bp mRNA linear EST 21-NOV-2000 DEFINITION 601888252F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4122276 5',	ACCESSION BF304699 VERSION BF304699.1 GI:11251586 KEYWORDS EST. SOURCE HOMO Sapiens (human)	ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalla; Eutherla; Primates; Catarrhini; Hominidae; Homo.	REFERENCE 1 (bases 1 to 94%) AUTHORS NIH-MGC http://mgc.ncl.nih.gov/. TITLE National Institutes of Health, Mammalian Gene Collection (MGC) JOURNAL Unpublished COMMENT Contact: Robert Strausberg, Ph.D. Emall: cgapbs-rémall.nih.gov Tissue Procurement: ATCC	cDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC Clone distribution information can be	found through the I.M.A.G.E. Consortium/LIML at: image.llnl.gov Plate: LLCMI005 row: g column: 13 High quality sequence stop: 646. FEATURES Location/Qualifiers Source /organism="Homo sapiens" //mol_type="MRNM" //mol_type="MRNM"	/////	Site_2: XhoI; cDNA made by oligo dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >SOUbp for average insert size 1.8bb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (Iniversity	of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)." BASE COUNT 133 a 329 c 351 g 171 t	Alignment Scores: 26.1 Length: 984
Oy 178 ValalaysalaValAspPherrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrrr	Qy 187 ValGluserLeuGlu 191 Db 193 TIGGATACATITGAG 179	RESULT 9 CC406705 LOCUS CC406705 CC406705 APP DNA linear GSS 19-MAY-2003 DEFINITION PUHKD12TD ZM.0.6_1.0_KB Zea mays genomic clone ZM:0F409B24,	NOI N DS	_		COMMENT Other_GSSs: PUHKD12TB COMMENT Other_GSSs: PUHKD12TB Contact: Cathy Whitelaw TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org	Class: sheared ends. FEATURES Location/Qualifiers source 1789	/organism="free mays" /mol_type="genomic DNA" /strain="813" /db_xref="taxon:4577" /clone="xmmra469B24" /clone="xmmra469B24" /clone="xmmra469B24" /clone="xmmra469B24" /clone="xmmra469B24" /clone="xmmra469B24"		Alignment Scores: Pred. No.: 95.50 Matches: 49 Score: 95.50 Matches: 49 Percent Similarity: 26.494 Mismatches: 67 Query Match: 9.404 Indels: 46 DB: 29 Gaps: 10	US-09-965-594-22 (1-197) x CC406705 (1-789) QY 29 GINGLYTHICLOLYSTHISTHISTHISTHSPAPAPASPLYSASGGIDVAlGLUGLYGLU 47 1	Oy 48 ValGlnIleValSerThrAlaThrGlnThrPheLevalaThrSerIleAsnGlyVal 66	Oy 67 LeutrpThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGlyProVal 86 ::: :::	Qy 87 ThrGlnMetTyrThrasnValaspLysAspLeuValGlyTrpGlnAlaPro 103

```
1062 bp DNA linear GSS 05-JUL-2001
T3 end of clone AW0AA006B03 of library AW0AA from strain CLIB 89 of
AL410673
AL410673.1 GI:12179275
end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in Ml3mpl9, .b/q reads in DQC19). The same ligation was transformed in either JM107 or DH5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           154 SerGlyGlyProLeu-LeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVa 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 90 TyrThrAsnValAspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArg---Ser 108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      LeuThrProCysThr----- 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         341 CTCCCTCCTGTGCTGTGCTCCACCCAGTGCCCCCGGCGCCGCGTACGGGGACAGGGA 282
                                                                                                                                                                                                                                                                                                                                       ....-.----AlaThrSerIleAsnGlyValLeuTrpThrValTyrHis 72
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              73 GlyAlaGlyThrArgThr-----IleAlaSerProLysGlyProValThrGlnMet 89
                                                                                                                                                                                                                                                                                                                   23 GlnGlnThrArgGlyGluGlnGlyThrGln---LysThrSerHisThrGlyArgAspLys 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Yarrowia lipolytica
Yarrowia lipolytica
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -----CysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 281 AACTGGTGTGGTGAGGATCACGTGGAGAGGCAGGACGACGACGACGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           116 CGGGGGGGGCCGTTGGATGTGCGTGGCAGTGGCAGCCGCT------GCGTACGT
                                                                                                                                                                                                                                                                                                                                                                                        42 AsnGlnValGluGlyGluValGlnIleValSerThrAlaThrGlnThrPheLeu----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           143 SerProArgProlleSerTyrLeuLysGly-------
                                                                                                                                                              701
61
17
70
70
69
                                                                                                                                                                             Matches:
Conservative:
Mismatches:
Indels:
                                                                                       110 t
                                                                                                                                                              Length:
                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                 US-09-965-594-22 (1-197) x BZ342381 (1-701)
                                                                                         232 g
                                                                                                                                                          18.8
95.00
36.11$
28.24$
9.35$
                                                                                       251 c
                                                                                                                                                                                             Percent Similarity:
Best Local Similarity:
                                                                                       108
                                                                                                                                             Scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   236
                                                                                                                                                                                                                                                                                                                                                                                                                           999
                                                                                                                                                                                                                                                                                                                                                                                                                                                             09
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             114
                                                                                                                                                                                                                               Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
CNSO6QHN
LOCUS
DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                Pred No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
                                                                                                                                           Alignment
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                      셤
                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /lab_host='JM107 or DH5a"
/clone_llb='WGS-SbicolorF (JM107 adapted methyl filtered)"
/clone_llb='WGS-SbicolorF (JM107 adapted methyl filtered)"
/clone_llb='WGS-SbicolorF (JM107 adapted wether was added by fill in the recessive 3' end. The genomic DNA was nebulized,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BE342381 100 CSS 06-NOV-2002 ic83b11.bl WGS-SbicolorF (JM107 adapted methyl filtered) Sorghum bicolor genomic clone ic83b11 5', genomic survey sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sorghum bicolor

Eukaryota; Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;

Eukaryota; Viridiplantae: Streptophyta; Embryophyta; Tracheophyta;

Eukaryota; Viridiplantae: Streptophyta; Embryophyta;

Clade; Pantcoideae; Andropogoneae; Sorghum.

I (bases 1 to 701)

Rabinowicz,P.D., O'Shaughnessy,A.L., Ballja,V., Dedhia,N.,

Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,

Zutavern,T., Palmer,L., McComble,W.R. and Martienssen,R.A.

Genomic shotgun sequences from Sorghum blcolor (methyl-filtered)
                                                                                                                                                                                                                                                                                                                                                                                                                                                554
                                                                                                                                                                                                                                                                                   139 GlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSerGlyGlyProLeu 158
                                                                                                                                                                                                                                                                                                                                                        159 LeuCysProAlaGlyHis-Ala-----ValGlyIlePheArg-----AlaAlaVa 173
                                                                                                                                                                                                                                                                                                                                                                                        526 TGGTGTCCAAGCTCGCGCTGCTGTACTGGAAGTCGGCACGCTCCGGTCAGGTGCAGCTTC 467
                                                                                                                                             100 TrpGlnAlaProGlnGlySerArgSerLeuThr...ProCysThrCysGlySerSerAsp 118
                                                                                                                                                                                                                                     553 GGGCGCCTCTTG-----TGGGGAAGACCTCGA
                                                                                                                                                                                                                                                                                                                                                                                                                           173 ISerThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrTh
                                                                                                                                                                                                              119 LeuTyrLeuValThrArgHisAlaAspVallleProValArgArgArgGlyAspSerArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Contact: W. Richard McCombie
Lita Annenberg Hazen Genome Sequencing Center
Cold Spring Harbor Laboratory
PO Box 100, Cold Spring Harbor, NY 11724, USA
Fax: 516 367 8884
   36
34
7
                    Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /organism="Sorghum bicolor"
/mol_type="genomic DNA"
/db_xref="taxon:4558"
/clone="ic83bll"
   Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Email: mccomble@cshl.org
Plate: ic83 row: b column: 11
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 701.
Location/Qualifiers
1. 701
                                                                          Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sorghum bicolor (sorghum)
                                                                                                          US-09-965-594-22 (1-197) x BF304699
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BZ342381.1 GI:24742983
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CAAGGGGGGCGCC 415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                rMetArgSerPro 197
95.50
42.868
37.148
9.408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unpublished
                    Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BZ342381
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GSS.
                                                                                                                                                                                                                                                  598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  427
                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
B2342381/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
```

a

a õ g ò 셤 ò QQ δ

ŏ 용 ò

Saccharomycetales; Dipodascacese: Yarrovia. 1 [bases 1 to 162] 2 [class 1 to 162] 3 sharing the state of the period of the pe	DD 525 TCCCAGGGATACTCTCAGACATGGGCCACCTACACCGATGAA 566 T HisGly	11 8 8 1 6 1 6 6 6 6 6 6 6 6 6 6 6 6 6 6	BM467279 BM467279 BM467279 BM467279 BM467279 DEFINITION AGENCOURT_6430725 NIH_MGC_85 Homo sapiens cDNA clone IMAGE.5505315 5', mRNA sequence. 5', mRNA sequence. BM467279 BM46	Tissue Procurement: Lou Staudt CDNA Library Preparation: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Place: LLAM12148 Town: Doclumn: 04 High quality sequence stop: 170. Location/Qualifiers J. 1440 /organism="Homo sapiens" /do_xref="taxon:9606" /clone="libe="wingma" /do_xref="taxon:9606" /clone="libe="wingma" /do_xref="taxon:9606" /clone="libe="wingma" /do_xref="taxon:9606" /clone="libe="wingma" /do_xref="bH10B (phage-resistant)" /clone="libe="wingma" /docation-libe="wingma" /docation-libe="wingm
9 A B C C C A A		\(\frac{1}{2}\)	LEGGE PER PER PER PER PER PER PER PER PER PE	T. 43
TITLE JOURNAL MEDLINE PUBMED AUTHORS AUTHORS AUTHORS AUTHORS AUTHORS TITLE JOURNAL MEDLINE PUBMED REFERENCE ROUN AUTHORS SOUR ALIGNMENT COMMENT COMMENT COMMENT Alignment Pred. No. SCOTE: Pred. No. SCOTE: Pred. No. SCOTE: Pred. No. COMMENT OUETY MAI	Saccharomycetales; Dipodascaceae: Yarrowia. 1 (bases 1 to 1062) Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrans,P., Lepingle,A., Llore Malpertuy,A., Neuveglise,C., Osler-Kalogeropoulos,O., Po Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louv Wincker,P. and Weissenbach,J. Genomic exploration of the hemiascomycetous yeasts: 1. A yeast species for molecular evolution studies FEBS Lett. 487 (1), 3-12 (2000) 20584711 1152876 2 (bases 1 to 1062) Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feyne Artiguenave,F., Wincker,P. and Galilardin,C. Genomic exploration of the hemiascomycetous yeasts: 17.	_	rce Lfe	1062 36.3 Longth: 1062 53 94.50 Matches: 53 53 Constructive: 53 53 Constructive: 53 53 Constructive: 54 54.20 Mismatches: 52 Constructive: 25 Mismatches: 62 Constructive: 25 Constructive: 25 Constructive: 29 Constructive: 29 Constructive: 29 Constructive: 29 Constructive: 29 Constructive: 29 Constructive: 20 Constructive

```
/dev_stage="adult"
/clone_lib="An expressed sequence tag (EST) collection
/clone_lib="An expressed sequence tag (EST) collection
/clone_lib="An expression plant Selaghinella lepidophylla"
/note="Vector: Lambda Uni.Zap XR. Bluescript SK:: Site_l:
EcoRi, Site_2: Xhol: Library construction was performed
according to manufacture's (Stratagene, Inc.) recommended
protocol for the Lambda UniZapXR vector and cDNA synthesis
kit.
                                                                                                                                                                                                                                                                                                                                                                                                                                /tlssue_type="microphyll fronds undergoing desiccation for 2.5 h"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94 AspLysAspLeuValGlyTrpGlnAlaProGlnGlySerArgSerLeuThrProCysThr 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     173 GGATTGGATCATACGCTGACCACAGGCGTCATCAGTGGTCTTCGAAGGGAGATTACT--- 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             151 LysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArg 170
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      171 AlaAlaValSerThrArgGlyValAlaLysAlaValAspPhe---IleProValGluSer 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 (bases 1 to 528)
Iturriaga,G. and Cushman,J.C.
An expressed sequence tag (EST) collection from the resurrection
plant Selaginella lepidophylla
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GlySerLeuLeuSerProArgProlleSerTyrLeu--------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GACAAGGATGTAGCGGTGCTGAAGATCGATGCTCAAGCAACAGATCTCAGCCCAATACCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               113 CITGGAAGTICGICCGATCIGCITGTIGGCCAGAAGGIGTATGCIATCGGTAAICCTTII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ------ThrArgHisAlaAspValIleProValArgArgArgGlyAspSerArg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    287 CCTGGGAACAGCGGGGGTCCGCTATTGGACAGTTCTGGAAATTTGATAGGCATCAACACT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               114 CysGlySerSerAspLeuTyrLeuVal------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             528
37
15
43
26
                                                                                                                                                                                                                                                                                                                                                                        /organism="Selaginella lepidophylla"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       137
                                                                                                                                                                                                                                                                                                                                                                                       /mol_type="mRNA"
/db_xref="taxon:59777"
/clone="SLA005F12"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-22 (1-197) x BM402566 (1-528)
                                                                                                                                                          \sigmaSA
                                                                                                                                                                                                                                                                                 column: 12
                                                                                                Contact: Cushman JC
Department of Blochemistry
University of Nevada
MS200, Reno, NV 89557-0014, U.
Tel: 715-784-1918
Fax: 775-784-1650
Email: jcushman@unr.edu
                                                                                                                                                                                                                                                                             Plate: 005 row: F column: 1
Seq primer: T3 20mer
High quality sequence stop: 5
Location/Qualiflers
1. 528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20.5
93.00
42.98%
30.58%
9.15%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       125 c
                                                                                                                                                                                                                                             FORWARD: T3 20mer
BACKWARD: T7 21mer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Leu 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   123
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                         source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
           REFERENCE
AUTHORS
TITLE
                                                                                  JOURNAL
                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                   COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BM402566 528 bp mRNA linear EST 01-JUL-2002 SAR005F12_34513 An expressed sequence tag (EST) collection from the resurrection plant Selagainella lepidophylla Selaginella lepidophylla cDNA clone SLA005F12 5, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1083 CCACACCCGCATTTTACTCCCCGCTATATTCTNTTATCTGNTGCGGNCGGACGCCCCC 1142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1143 TAGGGGTGGGGGGGGGGCCATTTTTCACGGTAGC --- ACATCGCCCCGCCCCCTCATTT 1199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Selaginella lepidophylla
Selaginella lepidophylla
Sukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Lycopodiophyta; Isoetopsida; Selaginellales; Selaginellaceae;
Selaginella.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          116
                                                                                                                                                                                                                                                                                                                                                                                                                                                             920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  980
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            117 -------SerAspLeuTyrLeuValThrArgHisAlaAspValIleProVa 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          149 rLeuLysGlySerSerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePh 169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               eArg-AlaAlaValSerThrArgGlyValAlaLySAlaValAspPheIleProValGluS 189
                                                                                                                                                                                                                                                                                                                 753 CAAATCACCATATCCGGAGATGTTGCTCTCTGT------TTTAAGCGTCACACA 800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  85 oValThrGlnMetTyrThrAsnValAspLysAspLeuValGlyTrp------ 100
                                                                                                                                                                                                                                                                                                                                                                                                                           89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               85
                                                                                                                                                                                                                                                                                                                                                     49
         full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

442 c 308 g 355 t 5 others
                                                                                                                                                                                                                                                                                                                                                                                                                                             ::: |||||||
861 CCTTCCCCCCAGGAAGCAAGATGTTTCTTTTTTTTTAGAAACGGGAGGCGTTGCTTC
                                                                                                                                                                                                                                                                               12 ArgileAsnLeuSerGlyAspThrAlaTyrAlaGlnGlnThrArgGlyGluGlnGlyThr
                                                                                                                                                                                                                                                                                                                                                     32 GlnLysThrSerHisThrGlyArgAspLysAsnGln-----ValGlu-GlyGluValGl
                                                                                                                                                                                                                                                                                                                                                                                   301 CCCGCGACCAACTCTACTGCCCGCCACAAAATATACTTTCGTGGAGCGGGAATATCTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                921 ATTITITITICCCIGGCAGGCACCTCTCGAATTCCCGAGCGCGCCGGTTCCCTCGCTCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----GlnAlaProGlnGlySerArgSerLeuThrProCysThrCysGlySer--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    131 lArgArgArgGlyAspSer-----ArgGlySerLeuLeuSerProArgProlleSerTy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           49 nIleValSerThrAlaThrGlnThrPheLeu---AlaThrSerIleAsnGlyValLeuTr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               68 pThrValTyrHisGlyAlaGlyThrArgThrIleAlaSerProLysGly-----Pr
                                                                                                                    1440
54
22
95
35
                                                                                                                    Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                        Indels:
                                                                                                                                                                                                                                               US-09-965-594-22 (1-197) x BM467279 (1-1440)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BM402566.1 GI:21643782
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1320 TICTCGTCACCACA 1333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   erLeuGluThrThr 193
                                                                                                                   54.2
94.50
36.89%
26.21%
9.30%
                                                                                                                                                                           Similarity:
                                                  ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BM402566
                                                  330
                                                                                                                                                        Percent Similarity:
                                                                                                  Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          101
                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
                                                                                                                                                                           Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
BM402566
                                           BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     음
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                셤
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                 셤
                                                                                                                                                                                                                                                                                                                                                                                   a
                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                     ò
```

346

```
129 eProValArgArgArgGlyAspSerArgGlySerLeuLeuSerProArgProlleSerTy 149
                                                                                                                                                                                                                                                                                                                                                                 149 rLeuLysGlySer-SerGlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIleP 169
            .----ArgSerLe 109
                                                                                                                          109 uThrProCysThrCysGlySerSerAspLeuTyrLeuValThrArgHisAlaAspValI1 129
                                                                                                                                                                                     -----CTGGAGCAGAT 277
                                                                                                                                                                                                                                                                                                         278 ATGACGAGCCATGGAGGGAGAGAGCAGGAGCCTTCTCCCTGGCCCAGCCCCTGGACCATÀ 337
                                                                                                                                                                                                                                                                                                                                                                                                                             338 TCTGAAAGCTGTGAAGTGGAGGC-----TGCCGGCGGGGGGGGCTGATGCAGCTGTTC 388
                                                              197 AGAGCTTGTGTGCTGGCAACACCCCCAGGGAGAACTTGCTGCATGGGGGAGGAGGTCTGT
      95 sAspLeuValGlyTrpGlnAlaProGlnGlySer-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Search completed: August 31, 2003, 04:27:50 Job time : 1914.31 secs
                                                                                                                                                                                  257 ACATAGATGT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        169 heArgAlaAla 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    389 TCACAGGGCT 399
                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                              d
                                                                                                                          ó
                                                                                                                                                                                                                                 οy
                                                                                                                                                                                                                                                                                                                                                              ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Other_GSSS: RPCI-II-32014.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 808 0200
F
                                                                                                                   AQS38021 550 bp DNA linear GSS 18-MAY-1999
RPCI-11-32014.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-32014,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     53 CACCCTGGAAGGCTATGAGATCATAGGAGGCTTCTGCCAGTCTCTGTCATCCACTGGC 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    113 CAAGAG------CACGGATTTGCAAAATGTGG------CACGGATTTGGA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            76 ThrArgThrIleAlaSerProLysGlyProValThrGlnMetTyrThrAsn-ValAspLy 95
                                                                                                                                                                                                                                                                                                                               Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 560)
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HisThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAlaThr 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56 GlnThrPheLeuAlaThrSerlleAsnGlyValLeuTrpThrValTyrHisGlyAlaGly 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Use of BAC End Sequences from Library RPCI-11 for Sequence Ready
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /db_xref="taxon:9606"
/db_xref="taxon:9606"
/cen="ReZ-1.11-32014"
/sex="Male"
/cen1_type="Lymphocytes"
/clone_lib="RPCI-11"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
a 125 c 181 g 124 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          560
44
111
54
36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7622691"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-22 (1-197) x AQ538021 (1-560)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                     genomic survey sequence
                                                                                                                                                                                                                                    AQ538021.1 GI:4849711
GSS.
                                                                                                                                                                                                                                                                                                      Homo sapiens (human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22.1
93.00
37.93%
30.34%
9.15%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Class: BAC ends.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Map Building
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ಸ
GTT 409
                                                                                                                                                                                                        AQ538021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          130
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
DB:
                                                                                                                      LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ..
9
                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          JOURNAL
COMMENT
                                                                                                                                                                                                              ACCESSION
                                                                                                                                                                                                                                          VERSION
KEYWORDS
SOURCE
                                                              RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                             AQ538021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
```

```
Hepatitis C virus
Hepatitis C virus
Hepatitis C virus
Synthetic nonstruc
Peptide KKNS4a use
                                                                                                                                         August 30, 2003, 17:42:58 ; Search time 2.49162 Seconds (without alignments) 700.745 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseqy-embl/AA1981.DAT:*
4: /SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*
5: /SIDSI/gcgdata/geneseqy-embl/AA1982.DAT:*
6: /SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseqy-embl/AA1985.DAT:*
8: /SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:*
9: /SIDSI/gcgdata/geneseqy-embl/AA1989.DAT:*
10: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
11: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
12: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
14: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
15: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
16: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
17: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
18: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
19: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
20: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
21: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
22: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
23: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
24: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
25: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
26: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
27: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
28: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
29: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
20: /SIDSI/gcgdata/geneseqy-embl/AA1991.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DAT: *
DAT: *
DAT: *
DAT: *
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /SIDS1/gcgdata/geneseq/geneseqp-emb1,AA2000.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1,AA2001.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1,AA2001.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1,AA2002.DAT:*/SIDS1/gcgdata/geneseq/geneseqp-emb1,AA2003.DAT:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Description
                                                                                                                                                                                                                                                                                                                                                                                                                                                              1107863
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                              1107863 segs, 158726573 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAB96864
AAW40552
AAM52536
AAG64158
AAB67391
AAB66371
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAB15227
AAM48242
                                                                                                using sw model
                                                                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A_Geneseq_19Jun03:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ΩI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                                                                                                                                                                US-09-965-594-26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query
Match Length DB
                                                                                                OM protein - protein search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23333333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1000.0
1000.0
1000.0
1000.0
1000.0
1000.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score
                                                                                                                                                                                                                                          Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                 Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     45000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Š
```

10							
11 49 100.0	10	49	0.00	23	23		peptide
13	_		0 00	23	23	AAE18687	
12	1.			9 6	3 6	AA1176276	District of the trees
14 49 100.0 23 24 AMB/5220 Hepatitis C VI	71.	,	0.00	57	7	AAU10370	ار
14 91 100 195 21 AAB15212 Hepatitis C vi	£.‡	5	0.00	73	24	ABG72264	U
15	14	6	0.001	195	21	AAB15212	
16	15	6	0.001	195	21	AAB15220	
17	16	6	0.001	200	13	AAR29846	
18	17	48	0.86	12	21	AAY44731	
19	18	48	0.86	13	22	AAB96862	
20 48 98.0 14 22 AAB74387 NS3 procease a AAB92237 48 98.0 14 22 AAB92237 Virus related 22 48 98.0 16 22 AAB9653 Hepatitis C Virus related 36 98.0 16 22 AAB96553 Hepatitis C Virus related 48 98.0 16 22 AAB96553 Hepatitis C Virus related 48 98.0 16 22 AAB96553 Hepatitis C Virus related 48 98.0 17 21 AAY83753 Hepatitis C Virus related 58.0 17 21 AAY83753 Hepatitis C Virus related 59.0 17 21 AAY83753 Hepatitis C Virus related 59.0 17 21 AAY83753 Hepatitis C Virus related 59.0 17 21 AAB96556 NS3 a period peptitis C Virus related 59.0 23 22 AAB96556 Hepatitis C Virus related 59.0 23 22 AAB96656 Hepatitis C Virus related 59.0 32 22 AAB96656 Hepatitis C Virus related 59.0 32 22 AAB96656 Hepatitis C Virus related 59.0 34 16 AAR82656 NS3 serine processor 54 8 98.0 54 19 AAW37808 NS4 proceed 70 AAW	19	4	0.86	7	18	AAW13792	Repatitis C virus
21 48 98.0 14 22 AAB92337 Virus related 23 48 98.0 16 22 AAB96851 Hepatitis C vi 24 48 98.0 16 22 AAB96853 Hepatitis C vi 25 48 98.0 16 22 AAB96854 Hepatitis C vi 26 48 98.0 17 21 AAY99552 27 48 98.0 17 21 AAY83773 HCV NS3A cofact 28 48 98.0 17 21 AAX83773 HCV NS3A cofact 29 48 98.0 17 21 AAX83775 HCV NS3A cofact 29 48 98.0 17 21 AAX83775 HCV NS3A cofact 29 48 98.0 17 21 AAX83775 HCV NS3A cofact 31 48 98.0 23 21 AAB92386 Hepatitis C vi 32 48 98.0 23 22 AAB96855 Hepatitis C vi 34 48 98.0 23 22 AAB96856 Hepatitis C vi 35 48 98.0 32 22 AAB96856 Hepatitis C vi 36 48 98.0 32 22 AAB96856 Hepatitis C vi 37 48 98.0 34 16 AAR82856 Hepatitis C vi 38 48 98.0 54 16 AAR82856 Hepatitis C vi 40 48 98.0 63 15 AAW90782 Hepatitis C vi 41 48 98.0 63 15 AAW90551 Hepatitis C vi 42 48 98.0 87 15 AAW90551 HCV peptide LIC 43 48 98.0 87 15 AAW90551 HCV peptide C vi 44 48 98.0 87 15 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 63 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 45 48 98.0 64 18 AAW90551 HCV peptide C vi 46 48 98.0 64 18 AAW90551 HCV peptide C vi 47 48 98.0 64 18 AAW90501 HCV VI 48 98.0 64 18 AAW	20	4 8	0.86	14	22	AAB74387	NS3 protease acti
22 48 98.0 16 21 AAPS5448 Peptitie 4AK us 24 48 98.0 16 22 AAB96853	21	4	98.0	14	22	AAB92337	Virus related pep
23 48 98.0 16 22 AAB96653 Hepatitis C vi	22	4	0.86	16	23	AAY54448	Peptide 4AK used t
24 48 98.0 16 22 AAB96853 Hepatitis C vi 25 48 98.0 16 22 AAB96854 Hepatitis C vi 26 48 98.0 17 21 AAY83773 Hepatitis C vi 26 48 98.0 17 21 AAY83773 HCV NS3A Cofface 48 98.0 17 21 AAY83773 HCV NS3A Cofface 28 48 98.0 17 21 AAY83775 HCV NS3A Cofface 28 48 98.0 17 21 AAY83775 HCV NS3A Cofface 28 48 98.0 17 21 AAB95367 NS4A peptide. 31 48 98.0 23 20 AAB96855 HCV Instruction of the peptide side of side of the peptide side of s	23	8	0.86	16	55	AAB96851	O
25 48 98.0 16 22 AAB96684 Hepatilis C vi 48 98.0 17 21 AAY89752 HCV NS3A Coffactor 48 98.0 17 21 AAY83775 HCV NS3A Coffactor 48 98.0 17 21 AAY83775 HCV NS3A Coffactor 48 98.0 17 21 AAY83775 HCV NS3A Coffactor 48 98.0 17 21 AAY87775 HCV NS3A Coffactor 48 98.0 17 21 AAY8775 HCV NS3A Coffactor 48 98.0 18 23 AB96855 Substrate peptide. 31 48 98.0 23 22 AAB96855 Virus related 34 48 98.0 23 22 AAB96855 Virus related 35 48 98.0 23 22 AAB96856 Hepatilis C vi 36 48 98.0 23 22 AAB96856 Hepatilis C vi 37 8erine pro 38 48 98.0 34 16 AAR82856 NS3 8erine pro 39 48 98.0 54 16 AAR82856 NS3 8erine pro 40 48 98.0 54 16 AAR82856 NS3 8erine pro 40 48 98.0 63 15 AAR49651 Hepatilis C vi 48 98.0 63 15 AAR49651 HCV peptide Cl 44 48 98.0 86 18 AAR49651 HCV peptide Cl 45 48 98.0 86 18 AAR49651 HCV peptide Cl 45 48 98.0 87 17 AAR95545 HCV I chimeri B1527 standard; protein; 11 AA. AAB15227 standard; protein; 11 AA. AAB15227 standard; protein; 11 AA.	24	8	0.86	16	22	AAB96853	U
26 48 98.0 17 21 AAY89552 Hepatilis C vi 27 48 98.0 17 21 AAY89373 HCV NS3A COfface 28 48 98.0 17 21 AAY83773 HCV NS3A COfface 29 48 98.0 17 22 AAB97114 Hepatilis C vi 30 48 98.0 18 23 20 AAY15763 Substrate peptide. 31 48 98.0 23 20 AAY15763 Substrate peptide. 32 48 98.0 23 22 AAB96655 Hepatilis C vi 34 48 98.0 23 22 AAB96655 Hepatilis C vi 35 48 98.0 23 22 AAB96656 Hepatilis C vi 36 48 98.0 34 16 AAR82866 Hepatilis C vi 37 48 98.0 34 16 AAR82866 Hepatilis C vi 40 48 98.0 54 19 AAW37808 Hopatilis C vi 41 48 98.0 63 15 AAW49651 Hepatilis C vi 42 48 98.0 63 15 AAW49651 HCV peptide CI 43 88 98.0 87 17 AAR9565 HCV peptide CI 44 48 98.0 87 17 AAR9565 HCV peptide CI 45 48 98.0 87 17 AAR9565 HCV peptide CI 46 98.0 87 17 AAR9565 HCV peptide CI 47 48 98.0 87 17 AAR9565 HCV peptide CI 48 98.0 87 17 AAR9565 HCV peptide CI 49 98.0 87 17 AAR9565 HCV peptide CI 40 98.0 87 17 AAR9565 HCV peptide CI 41 48 98.0 87 17 AAR9565 HCV peptide CI 42 48 98.0 87 17 AAR9565 HCV peptide CI 43 48 98.0 87 17 AAR9565 HCV peptide CI 44 48 98.0 87 17 AAR9565 HCV peptide CI 45 48 98.0 87 17 AAR9565 HCV peptide CI 46 98.0 87 17 AAR9565 HCV peptide CI 47 48 98.0 87 17 AAR9565 HCV peptide CI 48 98.0 87 17 AAR9565 HCV peptide CI 48 98.0 87 17 AAR9565 HCV peptide CI 49 98.0 87 17 AAR9565 HCV peptide CI 40 98.0 87 17 AAR9565 HCV peptide CI 40 98.0 87 17 AAR9665 HCV peptide CI 41 48 98.0 87 17 AAR9665 HCV peptide CI 42 48 98.0 87 17 AAR9665 HCV peptide CI 40 98.0 87 17 AAR9665 HCV peptide CI 40 98.0 87 18 AAR9665 HCV peptide CI 40 98.0 87 18 AAR9665 HCV peptide CI 41 48 98.0 87 18 AAR9665 HCV peptide CI 42 48 98 0 0 0 0 0 0 0 0 0 0 0 0 0 0	25	8	98.0	16	22	AAB96854	
27 48 98.0 17 21 AAY83773 HCV NS3A Cofac 28 48 98.0 17 21 AAY83775 HCV NS3A Cofac 29 48 98.0 17 21 AAB97114 Hepatitis C VI 30 48 98.0 18 23 AAB95367 Substrate peptide. 31 48 98.0 23 20 AAB95655 Hepatitis C VI 32 22 AAB96655 VITUAE Felated 33 48 98.0 23 22 AAB96655 VITUAE Felated 34 48 98.0 23 22 AAB96656 Hepatitis C VI 36 48 98.0 23 22 AAB96656 SISS SETIONE DEC 37 48 98.0 32 22 AAB96656 SISS SETIONE DEC 38 48 98.0 34 16 AAR82656 NS3 SETIONE DEC 39 48 98.0 35 19 AAW87386 HCV DEPTIGE USCALL 40 48 98.0 54 16 AAW87898 NATIONE DEC 41 48 98.0 63 15 AAW87651 HEPATITIS C VI 42 48 98.0 63 15 AAW89651 HCV DEPTIGE CI 43 48 98.0 63 15 AAW89651 HCV DEPTIGE CI 44 48 98.0 87 17 AAR95545 HCV DEPTIGE CI 45 48 98.0 87 17 AAR95545 HCV DEPTIGE CI 45 48 98.0 61 18 AAW89651 HCV DEPTIGE CI 45 48 98.0 87 17 AAR95545 HCV DEPTIGE CI 45 48 98.0 87 17 AAR95545 HCV DEPTIGE CI 45 48 98.0 87 17 AAR95527 HCV DEPTIGE CI 45 48 98.0 87 17 AAR95527 HCV DEPTIGE CI 45 48 98.0 87 17 AAR95527 HCV DEPTIGE CI 46 48 98.0 87 17 AAR95527 HCV DEPTIGE CI 47 48 98.0 87 17 AAR95527 HCV DEPTIGE CI 48 AAB15227 STANDARTS	56	4 8	0.86	17	21	AAY99552	tr)
28 48 98.0 17 21 AAY83775 HCV NSSA Cofac 29 48 98.0 17 22 AAAB97114 Hepatitis C vi 30 48 98.0 17 22 AAB97114 Hepatitis C vi 31 48 98.0 23 20 AAY15763 Substrate peptide. 32 48 98.0 23 21 AAB2336 Hepatitis C vi 34 48 98.0 23 22 AAB92336 Hepatitis C vi 35 48 98.0 28 12 AAB92336 Hepatitis C vi 36 48 98.0 32 12 AAB92336 Hepatitis C vi 37 48 98.0 34 16 AAR82856 NS3 serine produced by the serine produced by t	27	48	0.86	17	21	AAY83773	
29 48 98 0 17 22 AAB97114 Hepatliis C vi	28	8	0.86	17	21	AAY83775	HCV NS3A cofactor
30 48 98.0 18 23 ABB05367 NG4a peptide. 31 48 98.0 23 20 AAV15763 Substrate peptide. 32 48 98.0 23 20 AAB96855 Hepatitis C vide. 33 48 98.0 23 22 AAB96855 Hepatitis C vide. 34 48 98.0 23 22 AAB96856 Hepatitis C vide. 35 48 98.0 28 12 AAB96856 NG1aserine produced by the period of th	53	4.8	0.86	17	22	AAB97114	
31 48 98.0 23 20 AAY15763 Substrate pept 32 48 98.0 23 21 AAB2316 Substrate pept 33 48 98.0 23 21 AAB2316 Substrate pept 33 48 98.0 23 22 AAB2336 Hepatitis C vision 48 98.0 28 12 AAB2336 Hepatitis C vision 48 98.0 28 12 AAB3336 Hepatitis C vision 48 98.0 34 16 AAR82856 NS3 serine produced as 48 98.0 34 16 AAR82856 NS3 serine produced as 48 98.0 34 16 AAR82856 NS3 serine produced as 48 98.0 54 19 AAW37082 Peptide and 40 48 98.0 54 19 AAW37082 Hovy peptide cision 54 48 98.0 63 15 AAW37080 Native Hepatitis C vision 44 48 98.0 86 18 AAW30651 Hepatitis C vision 44 48 98.0 86 18 AAW30651 Hepatitis C vision 44 48 98.0 87 17 AAR95545 Hovy peptide cision 54 48 98.0 87 17 AAR95545 Hovy peptide Cision 54 54 98.0 87 17 AAR95545 Hovy peptide Cision 54 54 98.0 87 17 AAR95545 Hovy peptide Cision 54 54 98.0 87 17 AAR95545 Hovy peptide Cision 54 54 98.0 87 17 AAR95545 Hovy peptide Cision 54 54 98.0 87 17 AAR95545 Hovy peptide Cision 54 54 98.0 87 17 AAR95545 Hovy peptide Cision 54 54 98.0 87 17 AAR95545 Hovy peptide Cision 64 98.0 87 17 AAR95545 Hovy peptide Cision 54 98.0 87 17 AAR95545 Hovy peptide Cision 54 98.0 87 17 AAR95545 Hovy peptide Cision 64 98.0 87 17 AAR95545 Hovy peptide Cision 64 98.0 88.0 88.0 88.0 88.0 88.0 88.0 88.0	30	8	0.86	18	23	ABB05367	NS4a peptide. Hep
32 48 98.0 23 21 AAB22810 Synthetic pept; 34 48 98.0 23 22 AAB96855 Hepatitis C vii 34 48 98.0 23 22 AAB96855 Hepatitis C vii 35 48 98.0 23 22 AAB96855 Hepatitis C vii 35 48 98.0 28 19 AAB96856 Hepatitis C vii 37 48 98.0 34 16 AAR82865 Hepatitis C vii 38 19 AAR97082 Hepatitis C vii 38 19 AAR97082 Hepatitis C vii 38 19 AAR97082 Hepatitis C vii 48 98.0 34 16 AAR92855 Houstructural c vii 48 98.0 54 19 AAR97868 Hatte Hepatitis C vii 44 8 98.0 54 19 AAR97868 Hatte Hepatitis C vii 48 98.0 63 15 AAR9651 Hepatitis C vii 44 48 98.0 87 15 AAR49651 Hepatitis C vii 44 48 98.0 87 15 AAR49652 HCV peptide Cl4 48 98.0 87 17 AAR95845 HCV peptide Cl4 48 98.0 87 17 AAR95845 HCV II chimerid C vii 48 98.0 87 17 AAR95845 HCV II chimerid C vii 48 98.0 87 17 AAR95845 HCV II chimerid C vii 48 98.0 87 17 AAR95845 HCV II chimerid C vii 48 98.0 87 17 AAR95845 HCV II chimerid C vii 48 98.0 87 17 AAR95845 HCV II Chimerid C vii 48 98.0 87 17 AAR95845 HCV II Chimerid C vii 48 98.0 87 17 AAR95845 HCV II C vii 48 98.0 87 18 AAR95845 HCV II C vii 48 98.0 87 18 AAR95845 HCV II C vii 48 98.0 87 18 AAR95845 HCV II C vii 48 98.0 87 18 AAR95845 HCV II C vii 48 98.0 87 18 AAR95845 HCV II C vii 48 98.0 87 18 AAR95845 HCV II C vii 48 98.0 88 18 48 98.0 88 18 48 98.0 88 18 48 98.0 88 18 48 98.0 88 18 98 98 98 98 98 98 98 98 98 98 98 98 98	31	8	0.86	23	20	AAY15763	pept
33 48 98.0 23 22 AAB96855 Heparitis C vii 34 48 98.0 23 22 AAB96855 Virus related to 15 48 98.0 28 22 AAB96856 Heparitis C vii 36 48 98.0 32 22 AAB96856 Heparitis C vii 36 48 98.0 32 22 AAB96856 NS3 serios prof 37 48 98.0 34 16 AAR82856 NS3 serios prof 38 98.0 34 16 AAR82855 NS4A profein. 39 48 98.0 54 19 AAR82855 NS4A profein. 44 48 98.0 63 15 AAR82855 HCV peptide Cl4 44 48 98.0 63 15 AAR848651 Heparitis C L L L L L L L L L L L L L L L L L L	32	84	98.0	23	21	AAB23810	
34 48 98.0 23 22 AAB92336 Virus related in the partities C viring selection of the partities C viring selectin c viring selection of the partities C viring selection of the p	33	48	0.86	23	22	AAB96855	Hepatitis C virus
35 48 98.0 28 19 AAM97386 Hepatitis C vii 37 48 98.0 32 22 AAB98856 Hepatitis C vii 37 48 98.0 34 12 2 AAB98856 Hepatitis C vii 38 6 98.0 34 12 2 AAB98856 Hepatitis C vii 38 6 98.0 34 18 AAM97082 Peptide used in section of the sect	34	8	98.0	23	22	AAB92336	Virus related per
36 48 98.0 32 22 ARB96856 Heperitis C vi. 37 8810	35	8	0.86	28	19	AAW37386	Hepatitis C virus
37 48 98.0 34 16 AAR82856 NS3 serine prof 38 98.0 34 16 AAR82856 NS3 serine prof 38 98.0 36 19 AAAR82856 Peptide used in 39 48 98.0 54 19 AAAR82855 NS4A profeting 40 48 98.0 54 19 AAAR82855 NS4A profeting 41 48 98.0 54 19 AAAR82898 NS4A Profeting 42 48 98.0 63 15 AAAR8651 HCV Peptide Clid 44 48 98.0 86 18 AAAR8652 HCV peptide Clid 44 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 45 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 45 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 45 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 45 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 46 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 46 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 46 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 46 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 46 48 98.0 87 17 AAR865545 HCV II Chimeride Clid 47 AAB815227 standard; protein; 11 AA. AAB815227;	36	48	0.86	32	22	AAB96856	Hepatitis C virus
38 48 98.0 36 19 AAW50782 Peptide used in AAR82082	37	48	98.0	34	16	AAR82856	NS3 serine protea
39 48 98.0 54 16 AAR82855 NS44 Protein. 40 48 98.0 54 19 AAR82855 Nostructural 41 48 98.0 54 19 AAR82865 Nostructural 42 48 98.0 54 19 AAR82865 Nostructural 44 48 98.0 86 18 AAR86651 Hepatitis Clid 44 48 98.0 87 17 AAR89655 HCV peptide Clid 45 48 98.0 87 17 AAR89655 HCV peptide Clid 65 48 98.0 87 17 AAR89655 HCV peptide Clid 65 48 98.0 87 17 AAR89655 HCV II chimerid 81527 Standard; protein; 11 AA. AAB15227 standard; protein; 11 AA. AAB15227; AAB15227 standard; protein; 11 AA. AAB15227;	38	48	0.86	36	19	AAW50782	귝
40 48 98.0 54 19 AAW37808 Nonstructural of a 19 48 98.0 54 20 AAV37898 Nonstructural of a 2 48 98.0 54 20 AAV37898 Native HCV Nctive	33	48	0.86	24	16	AAR82855	
41 48 98.0 54 20 AAY17898 Native HCV NS4A 42 48 98.0 63 15 AAR49551 HCV peptide C14-43 48 98.0 86 18 AAR49652 HCV peptide C14-45 48 98.0 87 15 AAR49652 HCV peptide C14-45 48 98.0 87 17 AAR95545 HCV II chimeric C14-25 48 98.0 87 17 AAR95545 HCV II chimeric C14-25 48 98.0 87 17 AAR95545 HCV II chimeric C14-25 48 98.0 87 17 AAR95545 HCV II chimeric C14-25 48 98.0 87 17 AAR95545 HCV II chimeric C14-25 48 98.0 87 17 AAR95545 HCV II chimeric C14-25 48 98.0 87 17 AAR95527 Standard; protein; 11 AA. AAR95527 Standard; protein; 11 AA. AAR95527;	0	8	•	24	19	AAW37808	_
42	7	8	•	54	50	AAY17898	Native HCV NS4A
43 48 98.0 86 18 AANU9051 Hepatitis C VIVA 44 48 98.0 87 15 AAR95545 HCV II Chimeric 45 48 98.0 87 17 AAR95545 HCV II Chimeric ALIGNMENTS SULT 1 815.27 AAB15.27; AAB15.27; AAB15.27; 19-DEC-2000 (first entry)	42	8		93	15	AAR49651	HCV peptide C14-
44 48 98.0 87 15 ARR49652 HCV peptide C14- 45 48 98.0 87 17 ARR95545 HCV II chimeric ALIGNMENTS ALIGNMENTS AAB15227 AAB15227 AAB15227; 19-DEC-2000 (first entry)	.	4 8	•	86	18	AAW09051	
45 48 98.0 87 17 AAR95545 HCV II Chimeric ALIGNWENTS SULT 1 815227 standard; protein; 11 AA. AAB15227; AAB15227;		8		82	12	AAR49652	peptide
815227 815227 8AB15227 standard; protein; 11 8AB15227; 19-DEC-2000 (first entry)		₹		81	17	AAR95545	II
915227 AAB15227 standard; protein; 11 AAB15227; 19-DEC-2000 (first entry)							
SULT 1 815227 AAB15227 standard; protein; 11 AAB15227; 19-DEC-2000 (first entry)						ALIGNMENTS	
AAB1527 standard; protein; 11 AAB15227; 19-DEC-2000 (first entry)	RESULT 1						
AAB15227; 19-DEC-20	AMBIS22/ ID AMBIS	1227 St	andard;	prote			
19-DEC-2000		5227;					
19-DEC-2000							
		sc-2000		entr	γ)		

Hepatitis; NS3 protease; viral replication; chronic liver disease; liver failure; liver cancer. Wittekind M, Weinheimer S, Zhang Y, Goldfarb V; Hepatitis C virus NS4A residues 21-31 (BRIM) BRISTOL-MYERS SQUIBB CO. 06-JAN-2000; 2000WO-US00345. 99US-0115271. Hepatitis C virus WO200040707-A1. 08-JAN-1999; 13-JUL-2000.

Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C

WPI; 2000-465976/40.

KKNS4a peptide. S Hepatitis C virus Peptide KKNS4a for invention.

```
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              chimeric Hepatitis C virus (HCV)-pestivins genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a peptide from the tethered NS4A cofactor sequence from HCV NS3, which was used to illustrate the present
                                                              The prosent sequence comprises residues 21-31 of the Hepatitis C virus (HCV) NS4A protease enzyme. It was used to create a number of fusion proteins also containing the HCV NS3 protease. The NS4A and NS3 proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV) pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit HCV infection
                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present invention relates to a nucleic acid construct encoding a
                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                                                                                                                                               Score 49; DB 21; Length 11; Pred. No. 0.025; 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pestivirus; Npro; protease; NS3; screening; NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus NS4A cofactor peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAM48242 standard; Peptide; 12 AA.
                                  Claim 15; Page 22; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 1; Fig 2; 20pp; English.
                                                                                                                                                                                                                                                                                                                             100.0%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0344456.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hong Z, Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity 100.
                                                                                                                                                                                                                                                                                                                                                                                                  1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2002-121103/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                11 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                JS6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 34 - DEC - 2001
                                                                                                                                                                                                                                                                 complexes
                                                                                                                                                                                                                                                                                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAM48242;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                    ò
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The present invention describes methods and compositions capable of preventing the replication of hepatitis C virus (HCV), involving administering a compound which inhibits NS2/3 autocleavage. Also provided are peptides capable of inhibitting this cleavage step, of which this sequence is an example. These are useful in the treatment of HCV infection in humans and chimpenzees, and in research applications, for example in studying the stabilisation of NS2/3, the effects of NS2/3 on HCV polyprotein processing and the effects of inhibiting NS2/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cell, or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                 Gaps
                                                                                                                                                                                                                                                                                                             Hepatitis C virus; HCV; NS2; NS3; inhibitory peptide; cleavage; replication inhibition; chimpanzee; human; infection; gene therapy
                                                                                                                                                                                                                                                                                  Hepatitis C virus NS2/3 cleavage inhibitory peptide SEQ ID NO: 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Inhibiting hepatitis C virus (HCV) replication in HCV infected of a patient or treating a patient for HCV infection comprises inhibiting autocleavage of NS2/\beta .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ö
                                                                                 .;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 49; DB 22; Length 13; 100.0%; Pred. No. 0.03;
                                                    Length 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                100.0%; Score 49; DB 23; 100.0%; Pred. No. 0.027; ive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Disclosure; Page 46; 50pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAW40552 standard; peptide; 23 AA.
                                                                                                                                                                                                     AAB96864 standard; peptide; 13 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ÿ
                                                                                                                                                                                                                                                                                                                                                                                                                                    25-AUG-2000; 2000WO-US23444.
                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0151395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kuo
                                                                                                                                                                                                                                                        06-JUL-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11; Conservative
                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GSWIVGRIVE 11
                                                                                                        1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (MERI ) MERCK & CO INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Darke PL, Jacobs AR,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI; 2001-343059/36.
                                                  Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sest Local Similarity
                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         13 AA;
                         12 AA;
                                                                                                                                                                                                                                                                                                                                                                               WO200116379-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            30-AUG-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   autocleavage.
                                                                                                                                                                                                                                                                                                                                                                                                         08-MAR-2001
                                                                                11;
                                                                                                                                                                                                                              AAB96864;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                         Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches
                                                                                Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 4
AAW40552
ID AAW4
                                                                                                                                                                          RESULT 3
                                                                                                                                                                                     AAB96864
                                                                                                                                                                                                                  2 X 0
                                                                                                         ò
                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     a
```

Sun Aug 31 16:00:03 2003

AAW40552;

```
The present invention relates to a novel pyrrolopyrazinone derivative, its stereoisomer or salt. It was found that the derivative is a Hepatitis C virus (HCV) nonstructural (NS) 3 protease inhibitor. The derivative is useful for the manufacture of a medicament for the treatment of HCV and in therapy for treating HCV infection. The present peptide was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     The invention relates to novel ketoamide and ketoester derivatives for use as inhibitors of hepatitis C virus (HVV) NS3 protease inhibitors. The compounds are useful for treating viral infections e.g. hepatitis C virus. The present sequence was used in an experiment measuring the effect of an inhibitor on the rate of hydrolysis of an ester substrate.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New ketoamide derivatives useful for treating infections e.g. hepatitis
                                                                                                                                                                                      New pyrrolopyrazinone derivatives useful for treating Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus, HCV, NS3 protease, alpha-ketoamide inhibitor,
virucide, hepatotropic, antiinflammatory, viral infection, KKNS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 49; DB 22; 1
100.0%; Pred. No. 0.055;
.ive 0; Mismatches 0;
                                                                                                                                                                                                       infection are NS3 protease inhibitors -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Disclosure; Page 195; 282pp; English.
                                                                                                                                                                                                                                           Disclosure; Page 130; 191pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAG64158 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-DEC-2000; 2000WO-US32677
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         990S-0168998
                                     29-FEB-2000; 2000US-185618P.
28-FFB-2001; 2001WO-US06269
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (DUPO ) DU PONT PHARM CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 100.
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GSVVIVGRIVE
                                                                         DUPO ) DUPONT PHARM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 2001-464936/50.
                                                                                                                                                WPI; 2001-656752/75
                                                                                                                                                                                                                                                                                                                                                                                                                   23 AA;
                                                                                                            Zhang X, Han W;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W0200140262-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KKNS4a peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         03-DEC-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              07-JUN-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAG64158;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Han W;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAG64158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This is the amino acid sequence of the novel sNS4A (synthetic Hepatitis C nonstructural protein 4A) peptide. It acts as a cofactor for the NS3 protease in order to achieve proteclytic processing of Hepatitis C virus (HCV) nonstructural proteins. It is used in the method of the invention as part of a device which can be used to provide information for the design of drugs for the treatment of HCV infection. They can also be used for determining the 3-dimensional structure of molecules or molecular complexes which contain at least some structurally similar features to a HCV NS3 serine protease domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Virucide: pyrrolopyrazinone derivative; Hepatitis C virus inhibitor; nonstructural 3 protease; NS3 protease; viral infection.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New hepatitis C virus crystal compositions - comprising a HCV NS3-like polypeptide complexed with a NS4A-like polypeptide, used particularly for drug design
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 23;
                                                                                                        Synthetic Hepatitis C nonstructural protein; sNS4A peptide; cofactor; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100.0%; Score 49; DB 19; Length 2:
100.0%; Pred. No. 0.055;
wiematches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                     Thomson JA;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Morgenstern KA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Peptide KKNS4a used in an enzyme assay.
                                                                           Synthetic nonstructural peptide sNS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAM52536 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Claim 4; Page 30; 97pp; English.
                                                                                                                                                                                                                                                                                                       97WO-US16182.
                                                                                                                                                                                                                                                                                                                                        9605-0731336.
960S-0025274.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first entry)
                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local Similarity 100.
Marches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                  (VERT-) VERTEX PHARM INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                     Lin C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WPI; 1998-250953/22.
                                                                                                                                                                      Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                       Kim JL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO200164678-A2.
                                                                                                                                                                                                                           WO9811134-A1
                                                                                                                                                                                                                                                                                                     12-SEP-1997;
                                                                                                                                                                                                                                                                                                                                        18-OCT-1996;
12-SEP-1996;
                                     20-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31-JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  07-SEP-2001.
                                                                                                                                                                                                                                                                 19-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic
```

AAM52536;

RESULT 5

셤

õ

Sequence

FOX T,

; 0

Gaps

```
liver cancer; NS3; antiviral agent.
                                                                                                                                                                                                                                         23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200248116-A2.
                                 W0200102424-A2.
                 Unidentified
                                                                                 07-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-2002.
                                                11-JAN-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glunz PW,
                                                                                                                                                                                                                                                                                                                                                                 ABG32508;
                                                                                                                                                                                                                                          Seguence
                                                                                                                                                                                                                                                                                                                                RESULT 9
                                                                                                                                                                                                                                                                                                                                                          ద
ö
                                 ô
                                                                                                                                                                                                                                                                                                                                                               The present invention relates to Lactam derivatives. These derivatives may be used for treating hepatitis C virus (HCV) infection. They can also be used for inhibiting HCV in a body fluid sample and as a standard or reagent in a test or assay for determining the ability of a potential pharmaceutical to inhibit HCV NS3 protease and/or HCV growth,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus; protease; boronic acid; inhibitor; liver cirrhosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                        New lactam derivatives are hepatitis C virus NS3 protease inhibitors useful for treating HCV infections
                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                 ó;
                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 49; DB 22; Length 23; 100.0%; Pred. No. 0.055; ive 0; Mismatches 0; Indels
                100.0%; Score 49; DB 22; Length 23;
Larity 100.0%; Pred. No. 0.055;
Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus protease inhibitor related peptide #1.
                                                                                                                                                                          Lactam; hepatitis C virus; HCV; NS3 protease.
                                                                                                                                                                                                                                                                                                                                                 Example 26; Page 100; 130pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AAB66371 standard; peptide; 23 AA.
                                                                                                         AAB67391 standard; peptide; 23 AA.
                                                                                                                                                                                                                                          26-JUL-2000; 2000WO-US20189
                                                                                                                                                                                                                                                         99US-0145631
                                                                                                                                         26-APR-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 100.
                                                                                                                                                                                                                                                                         (DUPO ) DU PONT PHARM CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 GSWIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3 GSVVIVGRIVL 13
                                                 1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                         WPI; 2001-159696/16.
                         Local Similarity
nes 11; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                         23 AA;
 Š
                                                                                                                                                         KKNS4a peptide
                                                                                                                                                                                                         WO200107407-A1
                                                                                                                                                                                                                                                                                         Priestley ES,
23
                                                                                                                                                                                                                                                         26-JUL-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  09-APR-2001
                                                                                                                                                                                                                         01-FEB-2001
                                                                                                                                                                                          Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAB66371
                                                                                                                         AAB67391;
 Sequence
                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                           Best Loc
Matches
                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  XX DEX DIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                 g
```

```
The present invention provides a number of boronic acid derivatives which act as inhibitors of the hepatitis C virus NS3 protease enzyme. They can be used to treat infection by the virus, which can cause liver cirrhosis and liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New pyrimidinones useful as serine protease inhibitors in the treatment of e.g. viral infection \cdot
                                                                                                                                                                                                                                                                                                                                  New boronic acid derivatives, optionally containing peptides, used to treat hepatitis C infections, are hepatitis C viral protease inhibitors
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NS3: HCV; proteage; HCV infection; hepatitis; cirrhosis; liver cancer; pyrimidinone; serine proteage inhibitor; virucide; hepatotropic; antiinflammatory; blood plasma processing; KKNS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 100.0%; Score 49; DB 22; Length 23; Best Local Similarity 100.0%; Pred. No. 0.055; Matches 11; Conservative 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Peptide KKNS4a for HCV NS3 protease kinetic assay.
                                                                                                                                                                                               Forsyth TP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (BRIM ) BRISTOL-MYERS SQUIBB PHARMA CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 60; Page 208; 258pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ABG32508 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3
                                                                                                                                                                                               Kettner CA, Jagannathan S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12-DEC-2001; 2001WO-US47911
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13-DEC-2000; 2000US-255290P
07-JUL-2000; 2000WO-US18655.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-NOV-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 GSVVIVGRIVL 13
                                                                                                                               (DUPO ) DU PONT PHARM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Douty BD,
                                                                                                                                                                                                                                                              WPI; 2001-103001/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 2002-627251/67
```

```
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                         a
                                                                                                                                                                                                                 ö
                      The invention relates to pyrimidinones of a formula given in the claims of the specification, their stereoisomers, salts and prodrugs. In assays, the pyrimidinone compounds inhibited Hepatitis C virus (HCV) NS3 protease with IC_5 values of less than 100 micro M. The compounds are useful for treating viral infection e.g. HCV infection (the causative agent of acute hepatitis and associated with cirrhosis and liver cancer) and as a reagent used as inhibitors of HCV protease in the processing of blood plasma for diagnostic and other commercial purposes. The present sequence is a peptide, KNNS4a, used in an NS3 kinetic assay.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  salts or products which are useful as serine protease inhibitors.
The inidazolidinones of the invention may have virucide, hepatotropic, or antiinflammatory activities and may be used as a serine protease inhibitor (preferably Hepatitis C virus (HCV) NS3 protease inhibitor) or a HCV growth inhibitor. Compounds of the invention are useful for treating viral infection e.g. hepatitis C virus (HCV) infection and as a reagent used as inhibitors of HCV protease in the processing of blood plasma for diagnostic and other commercial purposes. The imidazolidinones of the invention inhibit HCV NS3 protease and/or HCV sequence represents the KNS4a peptide used in the blood plasma assay. The present experiments in the examples of the specification.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This invention relates to novel imidazolidinones or their stereoisomers,
                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New imidazolidinones useful as serine protease inhibitors in the treatment of e.g. viral infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV; hepatitis C; imidazolldinone; serine protease inhibitor; virucide; hepatotropic; antlinflammatory; NS3 protease; KKNS4a; growth inhibitor; viral infection; blood plasma processing.
                                                                                                                                                                                                                   ;
0
                                                                                                                                                                                      100.0%; Score 49; DB 23; Length 23; 100.0%; Pred. No. 0.055; ive 0; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (BRIM ) BRISTOL MYERS SQUIBB PHARMA CO.
Example 140; Page 192; 270pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Example 20; Page 112; 173pp; English.
                                                                                                                                                                                                                                                                                                                                         ABG31914 standard; Peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12-DEC-2001; 2001WO-US47916
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13-DEC-2000; 2000US-255168P
                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                      Local Similarity 100.
                                                                                                                                                                                                                                            1 GSVVIVGRIVE 11
                                                                                                                                                                                                                                                            3 GSWIVGRIVE 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2002-599498/64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ş
                                                                                                                                                               23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                         KKNS4a peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W0200248157-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23
                                                                                                                                                                                                                                                                                                                                                                                              05-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20-JUN-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Synthetic
                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                    ABG31914;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Han O;
                                                                                                                                                                                                         Best Loca
Matches
                                                                                                                                                                                                                                                                                                               RESULT 10
                                                                                                                                                                                                                                                                                                                           8 X 0 0 0 0 0 0 0 0 0 0 X 8
                                                                                                                                                                                                                                                                      a
                                                                                                                                                                                                                                               Š
```

```
ö
                                                 ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              The present invention relates to hepatitis C virus (HCV) core antigen and NS (nonstructural) 3/4a antibody combination assay that can defect both HCV antigens and antibodies present in a sample using a single soild matrix as well as immunoassay soild supports for use in the assay. The soild support is useful for detecting HCV infection in a boloogical sample. The present sequence is NS4A peptide which is used to purify NS3/4a conformational epitope in the exemplification of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ñ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Immunoassay solid support useful for detecting hepatitis C virus infection in a biological sample, comprises at least one of HCV anti-core antibody and HCV NS3/4a epitope, bound to the support
                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus; NS3/4a antibody; HCV infection; NS4A peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                               ..
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 23;
  Length 23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     George-Nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 49; DB 23; Length 2:
100.0%; Pred. No. 0.055;
+*** 0: Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                         NS4A peptide used to purify NS3/4a conformational epitope.
                                                 Indels
Ouery Match
100.0%; Score 49; DB 23;
Best Local Similarity 100.0%; Pred. No. 0.055;
Matches 11; Conservative 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tandeske L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Example 2; Page 50; 87pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAU76376 standard; Peptide; 23 AA.
                                                                                                                                                                                                                                                      AAE18687 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     14-JUN-2001; 2001WO-US19369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GSVVIVGRIVL 13
                                                                                            1 GSVVIVGRIVL 11
                                                                                                                                       3 GSVVIVGRIVL 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2002-179522/23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                23 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           W0200196875-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Unident1fled.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         08-MAY-2002
                                                                                                                                                                                                                                                                                                                                            17-MAY-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20-DEC-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chien DY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AAU76376;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Seguence
                                                                                                                                                                                                                                                                                                 AAE18687;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                            RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAU76376
                                                                                                                                                                                                                                   AAE18687
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8×2×5
```

```
The present invention relates to a new immunoassay solid support consisting essentially of at least one hepatitis C virus (HCV) NS3/4a conformational epitope and a multiple epitope fusion antigen (MEFA).

Bound to the support. The NS3/4a conformational epitope and/or sample from an HCV-infected individual. The immunoassay of the invention sample from an HCV-infected individual. The immunoassay of the invention sample. The method of the invention provides a sensitive, accurate diagnostic and prognostic tool to provide adequate patient care and to prevent transmission of HCV by blood and by blood products, or by personal contact. Use of NS3/4a conformational epitope in combination with MEFA, provides a sensitive and reliable method for detecting early HCV seroconversion. Use of MEFA has the aded advantages of decreasing masking problems, improving sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit surface area of substrate, and improving substrate. Detection accuracy is increased and the use of false results is reduced because of the identification and the use of false results is reduced because of the incidence of false results is reduced because of the increased and the use of highly immunogenic HCV antigens which are present during the early stages of HCV seroconversion The present amino acid sequence represents the non-structural protein NS4A peptide sequence. The peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                    Hepatitis C virus; HCV; NS3/4a conformational epitope; seroconversion; immunoassay solid support; multiple epitope fusion antigen; MEFA; non-structural protein; NS4A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Immunoassay solid support, useful for detecting hepatitis C virus infection in biological sample, comprises HCV NS3/4a conformational epitope and multiple epitope fusion antigen bound to the support -
                                                                                                                                                                                                                                                                                                                                                                                                                 Coit D;
                Hepatitis C virus (non-structural protein) NS4A peptide sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                 Tandeske L, George-nasciemento C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          100.0%; Score 49; DB 23; Length 23; 100.0%; Pred. No. 0.055;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABG72264 standard; peptide; 23 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 3; Page 48; 92pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                           15-JUN-2000; 2000US-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                       14-JUN-2001; 2001WO-US19156
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dest Match

Best Local Similarity 100...

Best Local 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-MAR-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                               Arcangel P,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2002-090228/12.
                                                                                                                                                                                                                                                                                                                                                                          (CHIR ) CHIRON CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23 AA;
                                                                                                                                     Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                    Medina-selby A;
                                                                                                                                                                            WO200196870-A2.
                                                                                                                                                                                                                  20-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABG72264;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                 Chien DY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   epitope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABG72264
õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

```
The present invention relates to immunoassays comprising Hepatitis C virus (HCV) NS3/4a conformational epitope and multiple epitope fusion antigen (MERA), bound to a solid support. The NS3/4a epitope fusion to the fusion antigen (MERA), and the fusion antigen (MERA), and the fusion and the fusion and the fusion and fusion are useful for detecting funduoassays and methods of the invention are useful for detecting support provides a sensitive and reliable method for detecting early HCV serconversion. The assays can detect HCV infection caused by any fusion genotypes of HCV. The use of the multiple epitope fusion proteins decreases masking problems, improves sensitivity in detecting antibodies by allowing a greater number of epitopes on a unit area of substrate, and improves selectivity. The present sequence representing HCV type I (HCV-1) NS4a peptide is used in a protease curyme activity assay in the examples of the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Immunoassay solid support for detecting Hepatitis C Virus infection in biological samples, comprises Hepatitis C Virus conformational epitope and multiple epitope fusion antigen -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coit D;
                                                                   fusion antigen;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   100.0%; Score 49; DB 24; Length 23; 100.0%; Pred. No. 0.055; ive 0; Mismatches 0; Indels
                                                 Immunoassay solid support; Hepatitis C Virus type-1; HCV-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   George-Nascimento C,
                                                                 NS3/4a conformational epitope; multiple epitope fusion
MEPA; anti-HCV antibody; NS3/4a conformational antigen;
HCV infection; E2 hypervariable region.
              Hepatitis C Virus type-1 (HCV-1) NS4a peptide.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tandeske L,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15212 standard; protein; 195 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Example 3; Page 17; 45pp; English.
                                                                                                                                                                                                                                                                       15-JUN-2000; 200QUS-212082P.
02-APR-2001; 2001US-280811P.
02-APR-2001; 2001US-280867P.
                                                                                                                                                                                                                                                                                                                                                                                                 GEORGE-NASCIMENTO C.
                                                                                                                                                                                                                                      14-JUN-2001; 2001US-0881654
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity 100.
Matches 11; Conservative
                                                                                                                                     Hepatitis C virus type 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Chien DY, Arcangel P,
Medina-Selby A;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                    (MEDI/) MEDINA-SELBY A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3 GSVVIVGRIVL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2003-147573/14.
                                                                                                                                                                                                                                                                                                                                            CHIEN D Y.
ARCANGEL P.
TANDESKE L.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23 AA;
                                                                                                                                                                      US2002146685-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                   COLT
                                                                                                                                                                                                        10-OCT-2002.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAB15212;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                 (TAND/)
(GEOR/)
(COIT/)
                                                                                                                                                                                                                                                                                                                                            CHIE/
                                                                                                                                                                                                                                                                                                                                                                ARCA/)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAB15212
ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0 % 2 % 5 % 5 %
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-MS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor
                                                                                                                                                                                                                                                                                                                                              Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                      The present sequence is a fusion protein created using the Hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; mutein.
                    Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100.0%; Score 49; DB 21; Length 195; 100.0%; Pred. No. 0.56; ive 0; Mismatches 0; Indels (
                                                                                                                                                                                                                                                                        Coldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus NS4A-NS3 fusion protease #2.
Hepatitis C virus NS4A-NS3 fusion protease #1.
                                                                                                                                                                                                                                                                      Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAB15220 standard; protein; 195 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                        Example 2; Fig 10; 66pp; English.
                                                                                                                                                                                                                                         (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                        Wittekind M, Weinheimer S,
                                                                                                                                                                                                            99US-0115271.
                                                                                                                                                                                06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 100.
Matches 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                    2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hepatitis C virus Synthetic.
                                                                           Hepatitis C virus
Synthetic.
                                                                                                                                                                                                                                                                                                                    N-PSDB; AAA73328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W0200040707-A1
                                                                                                                     WO200040707-A1.
                                                                                                                                                                                                            08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13-JUL-2000.
                                                                                                                                                   13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB15220;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a
```

```
The present sequence is a mutated version of a fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. This sequence contains the alpha-helix0-1
                                                                                                                                                                                                    Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              100.0%; Score 49; DB 21; Length 195; 100.0%; Pred. No. 0.56; tive 0; Mismatches 0; Indels
                                                                                                             Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    completed: August 30, 2003, 19:12:26
                                                                                                             Zhang Y,
                                                                          8
                                                                                                                                                                                                                                                                                                  Claim 23; Fig 12; 66pp; English.
                                                                        (BRIM ) BRISTOL-MYERS SQUIBB
                                                                                                             Wittekind M, Weinheimer S,
06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
N-PSDB; AAA73329.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 195 AA;
                                   08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Job time
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Search
a
```

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

- protein search, using sw model OM protein August 30, 2003, 19:02:22 ; Search time 0.905317 Seconds Run on:

(without alignments) 1168.492 Million cell updates/sec

US-09-965-594-26 49

1 GSVVIVGRIVL 11 Perfect score: Sequence: BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

283308 seqs, 96168682 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR_76:* Database

1: pirl:* 2: pir2:* 3: pir3:* 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ž

Description	qenome polyprotein		genome polyprotein				1	pro	genome polyprotein	genome polyprotein		polyprotein - hepa	probable copper-tr	probable dna polym	valine-tRNA ligase	H-2 class I histoc	. probable membrane	env polyprotein -	hypothetical prote	hypothetical prote	env polyprotein pr	env polyprotein pr							ت
DI 1	GNWVTC	GNWVCJ	A45573	S18030	GNWVTW	: PS0326	PC2219	-		-	GNWVC3			: T41395							VCLJEV	VCLJ22	VCLJEW	. VCLJE1	. VCLJE2	. VCLJE3	. VCLJWS	. VCLJE4	PS0130
Length DB	3010 1	3010 1	3010	3010 1	3010	492 2	876 2		_				8	**		44 2	ம		0	6	859 1	859	859 1	859 1	859	859	859]	860 1	37
Query Match	98.0	0.86	0.86	0.86	98.0	8.68	89.8	89.8	8.68	89.8	87.8	85.7	75.5	71.4	71.4	9 .69	4.69	69.4	† . 69	9 .69	69.4	69.4	69.4	4.69			69.4		67.3
Score	48	48	48	48	48	44	44	44	44	4	43	42	37	35	35	34	34	34	34	34	34	34	34	34	34	34	34	34	33
Result No.	-4	~	m	4	S	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	53

ą

Genome polyprotein - hepatitis C virus (strain J)
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructorein N:Contains: capsid protein C; envelope protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: hepatitis C virus
C;Accession: A3925; ps0086
R:Rato, N: Hijikata, M: Ootsuyama, Y: Nakagawa, M: Ohkoshi, S: Sugimura, T:: S}
Proc. Natl: Acad. Sci. U.S.A. 87, 9524-9528, 1990
A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patit

H-2 class I histoc H-2 class I histoc H-2 class I-like h	genome polyprotein genome polyprotein conserved hypothet	probable tail asse 3-oxoacyl-(acyl-ca conserved domain p conserved hypothet	hypothetical prote methenyltetrahydro lipoprotein (impor hymothaeleal prote	hypothetical prote argininosuccinate
PS0127 PS0118 A32934	PC1307 PC1306 D75611	D90908 AG3213 F95091 B97959	C85631 F69491 F89771 R71407	E90047 C83204
000	0000	7777	9999	100
37 38 39	142 209 220	248 271 291	300 316 331	395 405
67.3 67.3 67.3	67.3	67.3 67.3 67.3	67.3	67.3
		, w w w		. e. e.
30 31 32	₩ ₩ ₩ ₩ ★ ₩	3 8 8 8 8 9 8 9 9 9 9 9 9 9 9 9 9 9 9 9	4 4 4 4 0 11 9 6	44

AL IGNMENTS

```
genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 19-Jan-2001
                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                 ô
                                                                                                                                                                                                                                                                                                                          Score 48; DB 1; Length 3010;
Pred. No. 2.4;
                                                                                                                                                                                                                                                                                                                                               0; Indels
                                                                                                                                                                                                                                                                                                                                               1; Mismatches
                                                                                                                                                                                                                                                                                                                          98.0%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                    Best Local Similarity 90.9 Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                    1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                      ò
```

```
A.Note: this sequence is inconsistent with the nucleotide translation
A.Note: this sequence is inconsistent with the nucleotide translation
A.Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue ...
A.Note: sequence extracted from NBID backbone (NCBIN:121747, NCBIP:121748)
C.Superfamily: hepatitis C virus genome polyprotein
C.Seywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; St. F:2-115/Product: capsid protein C *status predicted <CPC>
F:16-191/Product: envelope protein M *status predicted <MED>
F:190-729/Product: nonstructural protein NSI *status predicted <NSI>
F:300-106/Product: nonstructural protein NSI *status predicted <NSI>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1317/Region: nucleotide-binding motif B (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Cispecies: hepatitis C virus
A; Variety: isolate JK1
C; Date: 19-May-2000 *text_change 23-Mar-2001
C; Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 23-Mar-2001
C; Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 23-Mar-2001
C; Accession: S18030; S33570; A48322; S18029
R; Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
Submitted to the FamEL Data Library, September 1991
A; Description: A whole genome of hepatitis C virus cDNA was isolated from a single p. A; Reference number: S18028
A; Reference number: S18030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         genome polyprotein – hepatitis C virus (strain Talwan)
N:Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: genomic RNA
A; Residues: 1-3010 <HON>
A; Residues: 1-3010 <HON>
A; Cross-references: BMBL: X61596; NID: 959478; PIDN: CAA43793.1; PID: 959479
A; Experimental source: isolate JK1 from an individual
B; Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A; Title: Sequence analysis of putative structural regions of hepatitis C virus isola*
A; Reference number: A48332; MUID: 93119270; PMID: 8380322
A; Accession: S33570
                                                                                                                                                                                                                                                                                                                                                                                                                     lenome polyprotein - hepatitis C virus (isolate JK1)
l;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F;1863-2013/Product: nonstructural protein NS4 #status predicted <N4B>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;204-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Molecule type: genomic RNA
A;Residues: 1-547,'T',549-621,'V',623-624,'S',626-652,'DL',655-761,'T',763-782 <HOW>
A;Cross-references: EMBL:X61591
                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 *sequence_revision 31-Dec-1992 *text_change 19-Jan-2001
                                                                                   Gaps
                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 48; DB 1; Length 3010;
       Length 3010;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                   Indels
       DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred. No. 2.4;
1; Mismatches
       Score 48; DB Pred. No. 2.4;
                                                                                   1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       98.08;
       98.04;
                                                                                                                                                                                              1678 GSVVIVGRIIL 1688
Query Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                       1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C; Accession: A40244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                             N; Contains:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 5
                                                                                                                                                                                                                                                                                                                                                                                                                         genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GNWVTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                       ò
                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A. Accession: A4573
A. Status: preliminary
A. McCession: A4573
A. Status: preliminary
A. Molecule (*Ppe: DAR)
A. Residues: 1-3010 crans
A. Residue: 1-3010 crans
A. Residue: 1-3010 crans
A. Residue: 1-3010 crans
C. Reywords: ATP: 9100protein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
C. Reywords: ATP: 9100protein; hydrolase; nucleotide Protein E *status predicted crass
F: 130-1306/Product: nonstructural protein NS1 *status predicted crass
F: 1007-1615/Product: nonstructural protein NS1 *status predicted crass
F: 130-1307/Region: nucleotide-binding motif A (P-loop)
F: 1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                          A;Accession: A39253
A;Molecule type: genomic RNA
A;Residues: 1-3010 <ARN-
A;Cross-references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
A;Cross-references: GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
B;Kato, N.; Ohkoshi, S.: Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A;Tille: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari
A;Reference number: PS0085
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome polyprotein - hepatitis C virus (strain JT)

W.Contains: capaba drotein C: envelope protein W: hepacivirin (EC 3.4.21.98) (nonstructury protein Ns4a: nonstructural protein NS4b: nonstructural protein NS4b: nonstructural protein NS5C (Species: hepatitis C virus
C: Species: hepatitis C virus
C: Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C: Accession: A45573
R: Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, A; Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: 8
A: Reference number: A45573; MUID: 92295714; PMID: 1318627
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  F;1616-1862/Product: nonstructural protein NS4a #status predicted <NA4>
F;1863-2013/Product: nonstructural protein NS4b fstatus predicted <NA4>
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98.0%; Score 48; DB 1; Length 3010; 90.9%; Pred. No. 2.4; Live 1; Mismatches 0; Indels
A; Reference number: A39253; MUID: 91088550; PMID: 2175903
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVL 11
```

a

```
A;Cross-references: GB:L29577; GB:L29578; GB:L29579
A;Experimental source: serum
C;Superfamily: hepatitis C virus genome polyprotein
                                                                                                                                                            C; Reywords: glycoprotein F:1-191/Product: core *status predicted <COE>F:68-78/Region: variable
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               89.8%;
81.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    89.8%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 81.8
Matches 9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1678 GCVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C; Species: hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Reference number: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 10; Conserv
                                       A; Residues: 1-876 <STU>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Accession: A36814
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GNWVCH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                              A; Recession: A40244
A; Molecule type: genomic RNA
A; Recidues: 1-3010 < CHED
A; Recidues: 1-3010 < CHED
A; Residues: 1-3010 < CHED
A; Cross-references: GB:NM4754
C; Superfamily: hepatitis C virus genome polyprotein
C; Reyords: ATP: capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
C; Reyords: ATP: capsid protein C *status predicted < CEPA>
F; 116-191/Product: envelope protein B *status predicted < NEE>
F; 130-729/Product: nonstructural protein NS1 *status predicted < NS1>
F; 130-1151/Product: nonstructural protein NS1 *status predicted < NS2>
F; 1310-11319/Region: nucleotide-binding motif A (P-loop)
F; 13110-1319/Region: nucleotide-binding motif B
F; 1316-1319/Region: nucleotide-binding motif B
F; 1316-1319/Region: nonstructural protein NS4 *status predicted < NAA>
F; 1863-2013/Product: nonstructural protein NS5 *status predicted < NAS>
F; 1014-3010/Product: nonstructural protein NS5 *status predicted < NS5>
F; 196, 209, 233, 234, 250, 305, 325, 417, 423, 430, 448, 532, 540, 556, 576, 623, 645, 1213, 1207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C.Accession: PS0326
Fili, U.S.; Tong, S.P.; Vitvitski, L.; Lepot, D.; Trepo, C.
Gene 105, 167-172, 1991
A.Title: Two French genotypes of hepatitis C virus: homology of the predominant genotype
A.Reference number: PS0326; MUID:92039028; PMID:1718820
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      polypeptide - hepatitis C virus (type 5a) (fragments)
N:Contains: core protein; El (carboxyl end); E2/NS1 (amino end); NS3 protein; NS4A prote
S:Species: hepatitis C virus
C:Date: 20-Feb-1995 *sequence_revision 20-Feb-1995 *text_change 17-Nov-2000
C;Accession: PC2219
R;Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S. Yirology 188, 102-113, 1992
A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the A;Reference number: A40244; MUID:92230206; PMID:1314449
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R;Stuyver, L.; Arnhem, W.V.; Wyseur, A.; Maertens, G.
Blochem. Biophys. Res. Commun. 202, 1308-1314, 1994
A;Title: Cloning and phylogenetic analysis of the core, E2, and NS3/NS4 regions of the
A;Reference number: PC2219; MUID:94338342; PMID:7520237
A;Accession: PC2219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        polyprotein - hepatitis C virus (isolate Fla) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 48; DB 1; Length 3010;
Pred. No. 2.4;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Accession: PS0326
A; Molecule type: genomic RNA
A; Residues: 1-492 <LIJ>
A; Cross-references: GB: M6020
A; Note: this sequence corresponds to nonstructural protein NS3
A; Note: translation of the nucleotide sequence is not complete
C; Superfamilly: hepatitis C virus genome polyprotein
C; Keywords: polyprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 44; DB 2; Length 492;
Pred. No. 2.3;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             89.8%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.08;
90.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match
Best Local Similarity 90.9'
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          229 GCVVIVGRIVI. 239
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        õ
```

```
A; Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 $sequence_revision 31-Dec-1992 $text_change 19-Jan-2001
C;Accession: A36814; A41546
R;Inchauspe, G:: Zebedee, S.; Lee, D.H.; Sugltani, M.; Nasoff, M.; Prince, A.M.
S;Inchauspe, G: Zebedee, S.; Lee, D.H.; Sugltani, M.; Nasoff, M.; Prince, A.M.
Submitted to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatitis C viru
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A Contents: annotation
A Note: neither amino acid nor nucleotide sequence is given
A Note: neither amino acid nor nucleotide sequence is given
C Superfamily: hepatitis C virus genome polyprotein
C Superfamily: hepatitis C virus genome polyprotein
C Steywords: ATP: capsid protein; envelope protein; glycoprotein; hydrolase; nonstruc
F:115/Product: capsid protein C *status predicted <CPC>
F:116-191/Product: envelope protein M *status predicted <CPC>
F:180-191/Product: major envelope protein MSI *status predicted <MED>
F:180-1006/Product: nonstructural protein NSI *status predicted <NSI>F:100-1016/Product: hepaclyinin *status predicted <NSI>F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1317/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: genomic RNA
A; Residues: 1-3011 - 1MC>
A; Residues: 1-3011 - 1MC>
A; Residues: 1-3011 - 1MC>
A; Ross-references: GB: MG7463; NID: 9329737; PIDN: AAA45534.1; PID: 9329738
A; Cross-references: GB: MG7463; NID: 92210296, 1991
A; Inchauspe, G: Zebedee, S: Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proco: Natl. Acad. Sci. U. SA. B8, 10292-10296, 1991
A; Title: Genomic structure of the human prototype strain H of hepatitis C virus: com A; Reference number: A41546; MUID: 92052256; PMID: 1658800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   fenome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstr
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
F:192-247/Product: E1 (carboxyl end) *status predicted <ERE>
F:248-411/Product: E2/NS1 (amino end) *status predicted <ENR>
F:248-318/Region: E2
F:348-411/Region: NS1 (amino end)
F:3412-783/Product: NS3 *status predicted <NSR>
F:3412-783/Product: NS4 *status predicted <NSR>
F:38-97/Product: NS4 *status predicted <NSR>
F:38-87/Product: NS48 *status predicted <NSR>
F:38-87/Product: NS48 *status predicted <NSR>
F:381,287,294,312,340/Binding site: carbobydrate (Asn) (covalent) *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 44; DB 1; Length 3011;
Pred. No. 13;
0; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 44; DB 2;
Pred. No. 4;
1; Mismatches
```

ö

Gaps

RESULT 9

```
A:Residues: 1577-1633 ACH2.
C:Superfamily: hepatitis C virus genome polyprotein
C:Reyvords: AFP: capsid protein: envelope protein: qlycoprotein; hydrolase; nonstruct
C:Reyvords: AFP: capsid protein C *status predicted <CPC.
F:116-191/Product: capsid protein M *status predicted <AEP.
F:190-729/Product: angor envelope protein B *status predicted <AEP.
F:300-729/Product: nonstructural protein NSI *status predicted <NSI>
F:107-106/Product: nonstructural protein NSI *status predicted <NSI>
F:107-107-1615/Product: hepacivitin *status predicted <NSI>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  genome polyprotein - hepatitis C virus (strain HCV-1)
N'Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
N'Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstru
C; Species: hepatitis C virus
C; Species: hepatitis C virus
C; Species: 30-69-1992 *sequence_revision 30-Sep-1992 *text_change 19-Jan-2001
C; Accession: A39166; PQ0403; PQ0404
R; Choo, O.L.; Richman, K.H.; Han, J.H.; Berger, K; Lee, C.; Dong, C.; Gallegos, C.;
Proc. Natl. Acad. Sci. US.A. 88, 2451-2455, 1991
A; Title: Genetic organization and diversity of the hepatitis C virus.
A; Reference number: A39166; MUID: 91172826; PMID: 1848704
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        F;1516-1862/Product: nonstructural protein NS4a *status predicted <NV4A>
F;1863-2013-Product: nonstructural protein NS4b *status predicted <NV8A>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NNS>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <NNS>
F;204-301/Product: nonstructural protein NS5 *status predicted <NNS>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A,Accession: A39166
A;Molecule type: mRNA
A;Cross-references: GB:M62321; NID:9329873; PIDN:AAA45676.1; PID:9329874
B;Chan, S.W.; McOmish, F.; Holmes, B.C.; Dow, B.; Peutherer, J.F.; Follett, E.; Yap, J. Gen. Virol. 73, 1131-1141, 1992
A;Title: Amalysis of a new hepatitis C virus type and its phylogenetic relationship A;Reference number: PQ0393; MUID:92268871; PMID:13156939
                                 F:731-1007/Product: nonstructural protein NSZ #status predicted <NS2>
F:1008-1616/Product: hepacivirin #status predicted <NS3>
F:1031-1338/Region: nucleotide-binding motif A (P-loop)
F:1317-1320/Region: nucleotide-binding motif B F:1317-1320/Region: nucleotide-binding motif B F:1317-1320/Region: nucleotide-binding motif F:183/Product: nonstructural protein NS4# *status predicted <N4A>
F:1864-2014/Product: nonstructural protein NS5# *status predicted <N4B>
F:2015-3014/Product: nonstructural protein NS5# status predicted <NS5>
F:2210-2249/Region: interferon sensitivity determining #status predicted
F;390-730/Product: nonstructural protein NS1 *status predicted <NS1>F;731-1007/Product: nonstructural protein NS2 *status predicted <NS2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         87.8%; Score 43; DB 1; Length 3011;
81.8%; Pred. No. 20;
iive 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              89.8%; Score 44; DB 1; Length 3014;
81.8%; Pred. No. 13;
Live 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A: Experimental source: isolates E-b16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Molecule type: genomic RNA
A;Residues: 1577-1633 <CHA>
A;Cross·references: DDBJ:D10128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1678 GCVVIVGRVVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             F:1316-1319/Region: DEXH motif
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity 81.8
Matches 9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 89.8
Best Local Similarity 81.8
Matches 9; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A:Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A:Accession: PQ0404
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A.Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
A.Experimental source: isolate HC-J1
C.Superfamilty: hepatitis C virus genome polyprotein
C.Steywords: ATP: glycoprotein: hydrolase; nucleotide binding; P-loop; polyprotein; p:116-191/Product: capsid protein C *status predicted <CPC>
F.116-191/Product: envelope protein M *status predicted <MED>
F.192-189/Product: major envelope protein E *status predicted <MED>
F.192-189/Product: nonstructural protein NS2 *status predicted <NS2>
F.1007-1615/Product: nonstructural protein NS2 *status predicted <NS2>
F.1230-1237/Region: nucleotide-binding motif B
F.1312-1317/Region: nucleotide-binding motif B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Experimental source: genotype 5a, which predominates in South Africa
A; Note: the translation of the nuclectide sequence is not complete in this paper
C; Superfamily: hepatitis C virus genome polyprotein
C; Superfamily: hepatitis C virus genome polyprotein
C; Keywords: ATP; glycoprotein: hydrolase; nuclectide binding; P-loop; polyprotein; E; 2-115/Product: capsid protein C *status predicted <CPC>
F; 16-19/Product: envelope protein B *status predicted <EPM>
F; 192-389/Product: anajor envelope protein E *status predicted <MEE>
F; 384-408/Region: hypervariable *status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructure)
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
Species: hepatitis C virus
C; Species: hepatitis C virus
C; Date: 19-May-2000 *sequence_revision 19-May-2000 *text_change 19-Jan-2001
C; Accession: JC5620
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Richamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.W.
Blocham. Blophys. Res. Commun. 236, 44-49, 1997
Rittle: The complete coding sequence of hepatitis C virus genotype 5a, the predominant A; Reference number: JC5620; MUID:97366593; PMID:9223423
A; Accession: JC5620
A; Molecule type: mRNA
A; Residues: 1-3014 <CHA>
A; Cross-references: GB:Y13184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Cross-references: EMEL: D10749; NID:q221586; PIDN:BAA01582.1; PID:q221587
R; Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990
A; Title: The S'-terminal sequence of the hepatitis C virus genome.
A; Reference number: PC1284; MUID:91013116; PMID:2170712
                                                                                                                                                                                     protein M; hepacivirin (EC 3.4.21.98) (nonstructu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                 genome polyprotein - hepatitis C virus
N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98
protein NG4s; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Species: 19-May-2000 #sequence_revision 19-May-2000 #text_change 19-Jan-2001
C;Accession: S40770; PC1285
R;Okamoto, H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                F;1316-1319/Region: DEXH motif
F;1316-1319/Region: DEXH motif
F;1616-1862-Product: nonstructural protein NS4a #status predicted <N4A>
F;2014-3011/Product: nonstructural protein NS5 *status predicted <N85>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       89.8%; Score 44; DB 1; Length 3011; 90.9%; Pred. No. 13; 1; Indels 1; of Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                             submitted to the EMBL Data Library, March 1992 A; Reference number: $40770 A; Accession: $40770
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1678 GCVVIVGRIVE 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local Similarity 90.9
es 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Molecule type: genomic RNA
A; Residues: 1-513 <OK2>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Molecule type: genomic RNA A; Residues: 1-3011 < OKA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
```

ö

Gaps

ö

Gaps

;

Length 574; Indels

```
C; Species: Staphylococcus aureus
C; Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C; Accession: B89949
R; Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; (
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, F
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Molecule type: DNA
A;Residues: 1-574 <000>
A;Cross-references: EMBL:AL023704; PIDN:CRA19261.1; GSPDB:GN00068; SPDB:SPCC553.09c
A;Experimental source: strain 972h-; cosmid c553
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-876 <RUN
A;Residues: 1-876 <RUN
A;Cross-references: GB:BA000018; PID:g13701460; PIDN:BAB42754.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: E89949
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 valine-tRNA ligase [imported] - Staphylococcus aureus (strain N315)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     71.4%; Score 35; DB 2; Length 876; 70.0%; Pred. No. 1.8e+02; tive 2; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                 Query Match 71.4%; Score 35; DB 2; 1 Best Local Similarity 77.8%; Pred. No. 1.2e+02; Matches 7; Conservative 2; Mismatches 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Search completed: August 30, 2003, 19:20:34 Job time : 2.90532 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C; Superfamily: valine-tRNA ligase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best_Local Similarity 70.0
Matches 7; Conservative
                                                                                                                                                                                                        A;Map position: 3
A;Introns: 89/2; 415/3; 518/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2 SWINGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          205 VVVGRIVV 213
                                                                                                                                                                                                                                                                                                                                                                                                                                 3 VVIVGRIVL 11
                                                                                                                                         C;Genetics:
A;Gene: SPDB:SPCC553.09c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                probable dna polymerase alpha-primase associated subunit - fission yeast (Schizosacchard Cispecies: Schizosaccharomyces pombe C;Species: Schizosaccharomyces pombe C;Date: 03-Dec-1999 *sequence_revision 03-Dec-1999 *text_change 03-Dec-1999 (C;Accession: T41395 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D. Babnited to the EMBL Data Library, May 1998 A;Reference number: 221991
                                                                                                                                                                                                                                                                                                                                                                                                                                 A.Molecule type: genomic RNA
A.Residues: 1-716 <KRE>
C.Superfamily: hepatitis c virus genome polyprotein
C.Superfamily: hepatitis; polyprotein
F:84,90,97,115,143,199,223,243,290,312/Binding site: carbohydrate (Asn) (covalent) *stat
                                                                                                                                                                                                                                   C;Accession: JQ1366
R;Kremsdorf, D.; Porchon, C.; Kim, J.P.; Reyes, G.R.; Brechot, C.
J Gen. Virol. 72, 2557-2561, 191
A;Title: A;Title: Toucheotide sequence analysis of a French hepatitis C virus: implication A;Reference number: JQ1366; MUIO:92013977; PMID:1655961
A;Accession: JQ1366
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ricole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Nathandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A; Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A; Title: Deciphering the blology of Mycobacterium tuberculosis from the complete genome A; Reference number: A70500; MUID:98295987; PMID:9634230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Realdues: 1.718 ccD.
A;Cross-references: GB:292771; GB:AL123456; NID:93242259; PIDN:CAB07083.1; PID:91877325
A;Experimental source: strain H37Rv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       C;Superfamily: Enterococcus copper-transporting ArPase cop8; ArPase nucleotide-binding F;128-454/Domain: ArPase transduction domain homology <ATT>
F;532-676/Domain: ArPase nucleotide-binding domain homology <ATN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    probable copper-transporting atpase 11/9 - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis C;Species: Mycobacterium tuberculosis C;Date: 17-Jul-1998 *sequence_revision 17-Jul-1998 *text_change 18-Aug-2000 C;Accession: G70978
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                      polyprotein - hepatitis C virus (French isolate) (fragments)
C:Species: hepatitis C virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   75.5%; Score 37; DB 2; Length 718; 63.6%; Pred. No. 63; Live 3; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 716,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 42; DB 2;
Pred. No. 7.6;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 85.7%;
Best Local Similarity 72.7%;
Matches 8; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity 63.6 nes 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

ö

Gaps

ő

```
GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
```

OM protein - protein search, using sw model

August 30, 2003, 18:01:52; Search time 0.544792 Seconds (without alignments) 949.524 Million cell updates/sec Run on:

US-09-965-594-26 49 1 GSVVIVGRIVL 11 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

127863 seqs, 47026705 residues Searched:

127863 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt_41:* Database : Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description		P26663 h genome po	c	c	h genome	æ	P26664 h genome po	mycobacte	074946 schizosacch	Q04487 saccharomyc			eduine		ednine				eduine C		Q9hy84 pseudomonas	-	_							-		8 lacto	90	P19324 mus musculu
SUMMARIES	Ę		POLG_HCVBK	POLG_HCVJA	POLG_HCVJT	POLG_HCVTW	POLG_HCVH	POLG_HCV1	CIPC_MYCTU	DPO2_SCHPO	YM07_YEAST	MCH_METTR	ENV_EIAV1	ENV_EIAV2	ENV_EIAV3	ENV_EIAV9	ENV_EIAVC	ENV_EIAVW	ENV_EIAVY	ENV_EIAV5	MCH_ARCFU	ASSY_PSEAE	SHU7_ECOLI	CYSG_ECOLI	JI60_HORVU	CTPC_MYCLE	MML2_MYCTU	- 1	COBM_RHOER	PLSX_THETN	Y8C8_SALTY		DEOB_LACIC	HS47_HUMAN	HS47_MOUSE
	Query Natch Length DR		3010 1			3010 1			718 1	574 1	1961	260 1	859 1	859 1	859 1	859 1	859 1	859 1	859 1	860 1	316 1	405 1	456 1	457 1	560 1	725 1	968 1	1158 1	249 1	332 1	388 1	411 1	411 1	417 1	417 1
	Query	Pa con	0.86	0.86	98.0	98.0	89.8	87.8	75.5	71.4	4.69	69.4	7 .69	69.4	69.4	69.4	69.4	69.4	69.4	4.69	67.3	67.3	67.3	67.3	67.3	67.3	67.3	67.3	65.3	65.3	65.3	65.3	65.3	vo.	65.3
	arc.	2000	48	48	48	4 8	44	43	37	35	34	34	34	34	34	34	34	34	34	34	33	33	33	33	33	33	33	33	32	32	32	32	32	32	32
	Result		7	7	٣	₹	Ŋ	9	7	80	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	5 4	25	26	27	28	29	30	31	• 32	33

P29457 rattus norv P50454 homo sanien	P44023 haemophilus Q10877 mycobacteri	P3629/ Saccharomyc Q8kcb7 chlorobium O45064 bacillus su	P03730 bacteriopha Q57140 haemophilus	P95993 sulfolobus Q8vi82 rattus norv	P38572 bos taurus
HS47_RAT CBP2 HUMAN	YFCC_HAEIN CTPB_MYCTU	1829_16831 RL27_CHLTE YNES BACSU	VTAI_LAMBD NEUA_HAEIN	RS2_SULSO FGFN_RAT	UPKA_BOVIN
	. – – .			~ ~	-
417	752	9 4 6 1 19 4 6 1	223	231 251	258
65.3	65.3	63.3	63.3	63.3 63.3	63.3
32	222	31	31	31	31
34	37	0 K 4	41	4 4 5 6 4	45

ALIGNMENTS

STANDARD; PRT; 3010 AA.	(Rel. 23, Crea (Rel. 23, Last	3 (Rel. 42, Last annotation update) yprotein (Contains: Capsid protein C (Core protein) (P22);	lycoprotein El (GP32) (GP35); Envelope glycoprotein E2 70) (NS1); Protein P7; Nonstructural protein NS2 (P21)	(EC 3.4.22): Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98): Nonstructural protein NS4A (P4): Nonstructural protein	(P27); Nonstructural protein NS5A (P56); Nonstructural protein (P66) (P70) (RNA-directed RNA polymerase) (EC 2.7.7.48)].	virus (isolate BK)	KNA POSILIVE-SLIANG VIIUSES, NO DNA	11105;	SEQUENCE FROM N.A. MEDLINE-91140698: PubMed-1847440:	Takamizawa A., Mori C., Fuke I., Manabe S., Murakami S., Fujita J.,	Unishi E., Andon T., roshida I., Okayama H.; "Structure and organization of the hepatitis C virus genome isolated	Carriers	Virol. 65:1105-1113(1991).	F 1487-1500.	Med-8647104;	C virus inhibits	phosphorylation mediated by cAMP-dependent protein kinase."; Bur J Biochem 237.611-618/1996)	OCINEMI: 237:011-010(1930).	X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1027-1215. MEDLINE-97015088: PubMed-RR61916:	Love R.A., Parge H.E., Wickersham J.A., Hostomsky Z., Habuka N.,	Moomaw E.W., Adachi T., Hostomska Z.; The crucks structure of bebattie C virue NG3 proteinses remosts a	a structural zinc binding site.";	1-342(1996).	X.FAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1027-1210 AND 1678-1691. MEDLINE-98227846; Pubmed-9568891;	Y., Munshi S., Sardana V., Cole J.L., Sardana M.,	er c., romer L., de Francesco K., Kuo L.C., chen Z.; If NS3 protease and NS4A peptide of BK strain hepatitis C	virus: a 2.2-A resolution structure in a hexagonal crystal form.";	FUNCTION: THE SMALL PROTEINS NS2A, NS2B, NS4A AND NS4B ARE	HYDROPHOBIC, SUGGESTING A POSSIBLE MEMBRANE-RELATED FUNCTION. NS3 AND NS5 MAY PLAY A ROLE IN THE VIRAL RNA REPLICATION.	ACTIVITY: Hydrolysis polyprotein, commonly	on, Cys or Thr in Pl and Ser or Ala in Pl'. TIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
T 1 HCVBK POLG_HCVBK	P26663; 01-AUG-1992 (Re 01-AUG-1992 (Re	15-SEP-2003 (Re Senome polyprot	Envelopė glycop (GP68) (GP70) ((EC 3.4.22); (EC 3.4.21.98);	NS4B (P27); Non NS5B (P66) (P70	itis	Hepacivirus.	NCB1_TaxID=11105; [1]	SEQUENCE FROM N.A. MEDLINE-91140698:	Takamizawa A.,	Unishi E., Ando "Structure and	from human carr	J. Virol. 65:11	SEQUENCE OF 1487-1500	MEDLINE-9623522 Rorowski P. He	Non-structural	phosphorylation m	[3]	X-RAY CRYSTALLO MEDLINE-9701508	Love R.A., Parg	MOOMBW E.W., Ad	trypsin-like fold and	Cell 87:331-342 [4]	X-RAY CRYSTALLO MEDLINE-9822784	Yan Y., Li Y.,	*Complex of NS3	virus: a 2.2-A	-!- FUNCTION: T	HYDROPHOBIC	-1- CATALYTIC A precursor p	position, C
30L 20L		DE				So		N N			¥ ¥		R L		RX		R.T		R X		RA PT		Z Z			¥			ខ្លួ	응응	ខខ

(POTENTIAL).

(POTENTIAL) (POTENTIAL) (POTENTIAL). (POTENTIAL). (POTENTIAL).

(POTENTIAL

a

```
PolG_HCVJA STANDARD; PRT; 3010 AA.
P26662;
01-AUG-1992 (Rel. 23, Created)
01-AUG-1993 (Rel. 41, Last sequence update)
28-FBB-2003 (Rel. 41, Last annotation update)
66-00me polyprotein (Contains: Capsid protein C (Core protein) (P22);
Envelope qlycoprotein El (GP32) (GP35); Envelope glycoprotein E2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 48; DB 1; Length 3010;
Pred. No. 1.1;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  327189 MW; F8422D5ECCFDFD9C CRC64;
                                                     GLCNAC
                  N-LINKED
N-L
                                                                                                                                                                                                                                                                                                                                                       LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match 98.0
Best Local Similarity 90.5
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                         10076
10085
10085
10085
10092
11093
11120
11120
11133
11134
11157
11168
11168
11168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1187 118
1189 1191
1198 120
1203 120
1680 168
                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                             CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
SARBOHYD
SATRAND
HELIX
STRAND
STRAND
STRAND
STRAND
STRAND
HELIX
TURN
TURN
TURN
HELIX
TURN
TURN
TURN
TURN
TURN
                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
STRAND
TURN
STRAND
STRAND
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 2
POLG_HCVJA
ID POLG_HC
AC P26662
DT 01-AUG.
DT 28-FEB
DT 28-FEB
DE GENOME
DE GENOME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HELIX
TURN
     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DR InterPro: IPR002512; HCV_capsid.

DR InterPro: IPR002512; HCV_capsid.

DR InterPro: IPR002513; HCV_core.

DR InterPro: IPR002513; HCV_core.

DR InterPro: IPR002519; HCV_NS2.

DR InterPro: IPR002519; HCV_NS3.

DR InterPro: IPR004109; HCV_NS4D.

DR InterPro: IPR004109; HCV_NS4D.

DR InterPro: IPR0071040; HCV_NS4D.

DR InterPro: IPR0071040; HCV_NS4D.

DR InterPro: IPR007094; RNA_pol_DS_PS.

DR InterPro: IPR007094; RNA_pol_DS_PS.

DR Pfan; PF01543; HCV_core: 1.

DR Pfan; PF01543; HCV_Core: 1.

DR Pfan; PF01549; HCV_NS1; 1.

DR Pfan; PF01560; HCV_NS1; 1.

DR Pfan; PF01560; HCV_NS1; 1.

DR Pfan; PF01000; HCV_NS3; 1.

DR Pfan; PF01560; HCV_NS3; 1.

DR Pfan; PF01640; HCV_NS4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELLULAR ANINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (CPOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
MAJOR ENVELOPE PROTEIN ES (POTENTIAL).
NONSTRUCTURAL PROTEIN NS1/E2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS2 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
POTENTIAL.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
CHARGE RELAY SYSTEM.
ATP (POTENTIAL).
                      ď
{RNA}(N).
-1- SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPOPROPIETH ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN W AND GLYCOPROTEIN B. THE NUCLEOCAPSID IS A COMPLEX OF
                                                                      PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                 EMBL; M58335; AAA72945.1; -
                                                                                                                                                                                                                                                                                 A38465; GNWVTC.
1A10; 25-MAR-98.
1JXP: 14-JAN-98.
1MS3; 08-APR-98.
1C2P; 15-NOV-00.
1CSJ; 08-NOV-99.
1GX5; 09-APR-02.
                                                                                                                                                                                                                                                                                                                                                                                                                        10UV; 26-JUN-00.
80HM; 20-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FRANSHEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                   PIR;
PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
PDB;
```

ö

Gaps

m

```
셤
                   ODDER SERVICE 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
(GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 34.22.-); Protease/helicase NS3 (P70) (Hepacivirin) (EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P67); Nonstructural protein NS5A (P86); Nonstructural protein NS5A (P86); Nonstructural protein NS5A (P86); Nonstructural protein Hepatitis C virus (isolate Japanese) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [RNA](N)
-1-SUBONT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDILINE-91088550; PubMed-2175903;
Kato N., Hilkata M., Ootsuyama Y., Nakagawa M., Ohkoshi S.,
Sugimura T., Shimotohno K.;
"Molecular cloning of the human hepatitis C virus genome from
"Appanese patients with non-A, non-B hepatitis";
Proc. Natl. Acad. Sci. U.S.A. 87:9524-9528(1990);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002166; HCV_RORP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_DOL_DS_PS.
INTERPO; IPR007094; RNA_DOL_PS_PS.
Pfan; PP01542; HCV_COP511; l.
Pfan; PP01542; HCV_COP511.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro: IPR001410; DEAD.
InterPro: IPR002522; HCV_capsid.
InterPro: IPR002521; HCV_core.
InterPro: IPR002531; HCV_core.
InterPro: IPR002519; HCV_NS2.
InterPro: IPR002518; HCV_NS2.
InterPro: IPR004109; HCV_NS3.
InterPro: IPR001490; HCV_NS3.
InterPro: IPR001490; HCV_NS4.
InterPro: IPR001490; HCV_NS4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; D90208; BAA14233.1; -. PIR; A39253; GNWVCJ. HSSP; P26663; 1JXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV_NS4a; 1
HCV_NS4b; 1
                                                                                                                                                                                                                                                    NCBI_TaxID-11116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4EROPS; S29.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PF01001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
      DESTRUCTION OF STREET S
```

```
ö
Pfam; PF01506; HCV_NS5a; 1.

Pfam; PF00271; helicase_C; 1.

Pfam; PF00998; Viral_RRMP; 1.

Probom; PD186062; HCV_NS1; 1.

R Probom; PD186062; HCV_NS1; 1.

R SWART; SM00487; DEXDc; 1.

R SWART; SM00487; DEXDc; 1.

R SWART; SM00487; DEXDc; 1.

R Core protein; Grat protein; Hydrolase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease.

T CHAIN 116 191 HARDEN CAPEID PROTEIN C BY THE CALL NOSTRUCTURAL PROTEIN C (POTEWTIAL).

T CHAIN 116 191 HAJOR ENVELORP PROTEIN E (POTEWTIAL).

T CHAIN 1100 NON-STRUCTURAL PROTEIN NS2 (POTEWTIAL).

T CHAIN 1101 1105 NON-STRUCTURAL PROTEIN NS2 (POTEWTIAL).

T CHAIN 1106 NON-STRUCTURAL PROTEIN NS4 (POTEWTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-GPG9) (GPG9) (Rel. 42, Last annotation update)
Envelope glycoprotein (GP32) (GP35); Envelope glycoprotein E2 (GP69) (GF09) (GP10) (NS1); Protein P7; Nonstructural protein NS2 (P21) (EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P50); Nonstructural protein NS5A (P60); CP7); Nonstructural protein NS5A (P60) (P70) (SNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RELAY SYSTEM (BY SIMILARITY).
RELAY SYSTEM (BY SIMILARITY).
RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLCNAC. . . . (POTENTIAL).
GGLCNAC. . . (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 48; DB 1; Length 3010;
Pred. No. 1.1;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MW; AA993794F46DB185 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC...) (IN-LINKED (GLCNAC...) (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
CHARGE RELA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHARGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 327017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.0%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local Similarity 90.5
Best Local Similarity 10.5
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2240
2788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ž
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-31642;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLG_HCVJT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FRANSMEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
POLG_HCVJT
```

```
NCBI_TaxID-31645;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLG_HCVTW
                                                                                                                                                                                                                                                                                              CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
CARBOHYD
                                                                                              ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                            TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P29846;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 4
POLG_HCVTW
                                                           CHAIN
     a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-slb.ch/announce/or send an email to license@isb-slb.ch).
         Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Tanaka T., Kato N., Nakagawa M., Ootsuyama Y., Cho M.J.,
Tanaka T., Kijikata M., Ishimura Y., Shimotohno K.;
"Molecular cloning of hepatitis C virus genome from a single Japanese
Tariter: sequence variation within the same individual and among
the virus Res. 23:39:34(192).

- Infected individuals.";
- Infected individuals...;
- Infected individuals...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CELLULAR AMINOPEPTIDASE.
CAPSID PROTEIN C (POTENTIAL).
MATRIX PROTEIN (POTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTRUCTURAL PROTEIN NSI,E2 (POTENTIAL).
NON-STRUCTURAL PROTEIN NS2 (POTENTIAL).
                                                                                                                                                                                                                                                                                        -!- SUBUNT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROPEIN HAND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                               -! - SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY 529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR001410; DEAD.
InterPro; IPR001221; HVV_capsid.
InterPro; IPR002512; HVV_capsid.
InterPro; IPR002513; HVV_capsid.
InterPro; IPR002519; HVV_core.
InterPro; IPR002519; HVV_NS1.
InterPro; IPR001409; HVV_NS2.
InterPro; IPR001409; HVV_NS3.
InterPro; IPR001409; HVV_NS4.
InterPro; IPR001409; HVV_NS4.
InterPro; IPR001409; HVV_NS4.
InterPro; IPR001045; HVV_NS4.
InterPro; IPR001066; HVV_NS4.
InterPro; IPR001094; NNA_pol_PSvir.
Pfam; PP01543; HVV_capsid; 1.
Pfam; PP01540; HVV_Capsid; 1.
Pfam; PP01550; HVV_LNS2; 1.
Pfam; PP0106; HVV_NS2; 1.
Pfam; PP0106; HVV_NS3; 1.
Pfam; PP0106; HVV_NS3; 1.
Pfam; PP0106; HVV_NS3; 1.
Pfam; PP0106; HVV_NS4; 1.
Pfam; PP0106; HVV_NS3; 1.
Pfam; PP0106; HVV_NS4; 1.
Pfam; PP0101; HVV_NS4; 1.
Pfam; PP0101; HVV_NS4; 1.
Pfam; PP0101; HVV_NS4; 1.
Pfam; PP01006; HVV_NS4; 1.
Pfam; PP01006; HVV_NS4; 1.
Pfam; PP01006; HVV_NS4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; D11168; BAA01943.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PIR: A45573; A45573.
PDB: 1A10; 25-WAR-98.
PDB: 1JXP; 14-JAN-98.
MEKOPS: 529.001; --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            115
191
383
729
1006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1116
192
384
730
                                                                                                                                                                                                                                                                            {RNA}(N).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       *** PEDLINE-92230206; PubMed-1314449;

*** MEDLINE-92230206; PubMed-1314449;

**A HEDLINE-92230206; PubMed-1314449;

**A HEDLINE-92230206; PubMed-1314449;

**Chen P.J., Lin M.H., Tal K.F., Liu P.C., Lin C.J., Chen D.S.;

**Chen P.J., Lin M.H., Tal K.F., Liu P.C., Lin C.J., Chen D.S.;

**The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' The The The Taiwanese hepatitis C virus genome: sequence determination and antigenomic RNA.*;

**L' Th
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-APR'1993 (Rel. 25, Created)
01-APR'1993 (Rel. 25, Last sequence update)
15-SEP-2003 (Rel. 25, Last sequence update)
66-00me polyprotein [Contains Capsid protein C (Core protein) (P22);
Envelope 91ycoprotein E1 (GP32) (GP35); Envelope 91ycoprotein E2 (GP68) (GP70) (MS1); Protein P7: Nonstructural protein NS2 (P21)
(EC 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
(EC 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS4B (P27); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); Constructural protein NS5A (P60) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
Hepatitis C virus (Isolate Taiwan) (HCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               the viral
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4A (POTENTIAL).
NONSTRUCTURAL PROTEIN NS4B (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                              SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           o,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 48; DB 1; Length 3010;
Pred. No. 1.1;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NKED (GLCNAC. . .) (POTI 94A1C77435D642BB CRC64;
                                                                                                                                                                                          CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
CHARGE RELAY SYSTEM (BY
                                                                                                                                                                                                                                                                                                                                                                                                                         N-LINKED GECNAC.
                                                                                                                                                                                                                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC
                                                                                                                                                                 POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  326573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ouery Match
Best Local Similarity 90.7%,
The Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1678 GSVVIVGRIIL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2041 204
2077 207
2240 224
2529 255
2788 278
```

us-09-965-594-26.rsp

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMEL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CELLULAR AMINOPEPTIDASE.
COELULAR AMINOPEPTIDASE.
CORE PROTEIN (POTENTIAL).
MATRIX PROTEIN (FOTENTIAL).
MAJOR ENVELOPE PROTEIN E (POTENTIAL).
NONSTROCTURAL PROTEIN NS.1 (POTENTIAL).
PROTEASE/HELICASE NS3 (POTENTIAL).
NONSTROCTURAL PROTEIN NS.4 (POTENTIAL).
NONSTROCTURAL PROTEIN NS.48 (POTENTIAL).
NONSTROCTURAL PROTEIN NS.48 (POTENTIAL).
RNA-DIRECTED RNA POLYMERASE (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Probom; PD186062; HCV_NS1; 1.
SMART; SM00487; DEXDc; 1.
Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                 SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY LIPPOPROTEIN BENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C. AND MRMA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REMOVED FROM CAPSID PROTEIN C BY THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                               SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC. . . .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (N-LINKED (N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro: IPR001410; DEAD.
InterPro: IPR001213; HCV_corp.
InterPro: IPR001221; HCV_corp.
InterPro: IPR001221; HCV_corp.
InterPro: IPR001219; HCV_env.
InterPro: IPR001219; HCV_NS1.
InterPro: IPR001219; HCV_NS3.
InterPro: IPR001409; HCV_NS3.
InterPro: IPR001409; HCV_NS4.
InterPro: IPR001409; HCV_NS4.
InterPro: IPR001409; HCV_NS5.
InterPro: IPR001409; HCV_NS5.
InterPro: IPR001409; HCV_NS5.
InterPro: IPR001509; HCV_NS5.
InterPro: IPR001509; HCV_NS1.
IPfam; PP01543; HCV_corp.
IPfam; PP01543; HCV_corp.
IPfam; PP01560; HCV_NS1; IPFam; PP01560; HCV_NS2; IPFam; PP01001; HCV_NS2; IPFam; PP01001; HCV_NS4; IPFam; PP01001; HCV_NS4; IPFam; PP01001; HCV_NS4; IPFam; PP01001; HCV_NS4; IPFam; PP01005; HCV_NS4; IPFam; PP01001; HCV_NS4; IPFam; PP01001; HCV_NS4; IPFam; PP01001; HCV_NS4; IPFam; PP01001; HCV_NS4; IPFam; PP001001; HCV_N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M84754; -; NOT_ANNOTATED_CDS.
PIR; A40244; GNWVTW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        729
1006
1006
2013
2010
3010
3010
1100
1116
1319
1319
1319
233
250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1N64; 25-FEB-03,
1NS3; 08-APR-98,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     347
1103
11165
11165
11165
1230
1233
233
233
250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEROPS; S29.001
MEROPS; U39.001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3D-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CHAIN
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                               X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1207-1657.
MEDLINE-9731322; Pubhed-9187654;
Yao N., Hesson T., Cable M., Hong Z., Kwong A.D., Le H.V., Weber P.C.;
"Structure of the hepatitis C virus RNA helicase domain.";
Nat. Struct. Biol. 4:463-467(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Kim J.L., Morgenstern R.A., Griffith J.P., Dwyer M.D., Thomson J.A.,
Murcko M.A., Lin C., Caron P.R.;
"Hepatitis C virus NS3 RNA helicase domain with a bound
oligonucleotide: the crystal structure provides insights into the mode
of unwinding.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Structure 6:89'100(1998).

-I- FUNCTION: PROTEASE NS2 IS RESPONSIBLE FOR THE CLEAVAGE OF NS2-NS3.

-I- FUNCTION: PROTEASE NS3 IS RESPONSIBLE FOR THE CLEAVAGE OF NS3-NS4A, NS4A-NS4B, NS4B-NS5A AND NS5A-NS5B.

-I- FUNCTION: NS4A FORMS A COMPLEX WITH NS3 AND IS ESSENTIAL FOR THE ACTIVATION OF NS3.
                                                                                                                                                                                                                                                                Gaps
305 N-LINKED (GLCNAC. ...) (POTENTIAL).
423 N-LINKED (GLCNAC. ...) (POTENTIAL).
429 N-LINKED (GLCNAC. ...) (POTENTIAL).
448 N-LINKED (GLCNAC. ...) (POTENTIAL).
448 N-LINKED (GLCNAC. ...) (POTENTIAL).
532 N-LINKED (GLCNAC. ...) (POTENTIAL).
556 N-LINKED (GLCNAC. ...) (POTENTIAL).
556 N-LINKED (GLCNAC. ...) (POTENTIAL).
557 N-LINKED (GLCNAC. ...) (POTENTIAL).
645 N-LINKED (GLCNAC. ...) (POTENTIAL).
527 N-LINKED (GLCNAC. ...) (POTENTIAL).
528 N-LINKED (GLCNAC. ...) (POTENTIAL).
5282 N-LINKED (GLCNAC. ...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Genomic structure of the human prototype strain H of hepatitis C virus: comparison with American and Japanese isolates."; Proc. Natl. Acad. Sci. U.S.A. 88:10292-10296(1991).
                                         (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-92052256; PubMed-1658800;
Inchauspe G., Zebedee S., Lee D.H.H., Sugitani M., Nasoff M.,
                                                                                                                                                                                                                                  Score 48; DB 1; Length 3010;
Pred. No. 1.1;
1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1192-1657.
MEDLINE-98154321; PubMed-9493270;
                                                                                                                                                                                                                                                                                                                                                                           PRT; 3011 AA
                                                                                                                                                                                                                                   98.08;
                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
  305
4417
4430
4430
5532
5540
576
623
6623
2041
2077
20529
2788
                                                                                                                                                                                                                                                                                     1 GSVVIVGRIVE 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11108;
  Prince A.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                         POLG_HCVH
P27958;
                                       CARBOHYD
                                                                CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                 CARBOHYD
CARBOHYD
                                                                                                        CARBOHYD
                                                                                                                                             CARBOHYD
                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                         SEQUENCE
                CARBOHYD
                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                 RESULT S
POLG_HCVH
                                                                                                                                                                                                                                                                                                                                                                          ô
                                                                                                                                                                                                                                                                                                            q
```

```
SIMILARITY).
SIMILARITY).
SIMILARITY).
                                                                                                                                                      (POTENTIAL).
     NONSTRUCTURAL PROTEIN NSSA.
                                                             (BY
(BY
                                                         CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
CHARGE RELAY SYSTEM
                                                                                                                                                    N-LINED (GLCNAC,
N-LINKED GLCNAC,
N-LINK
                                                                                                                  ATP (POTENTIAL).
DECH BOX.
                                           POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
N-LINKED
                                                                                                                                                                                                                                                   1319
1335
1340
1347
1353
1361
1368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1385
1393
1409
1411
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1463
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1480
1488
1490
1501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1317
1323
1336
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1343
1352
1362
1368
1373
1376
1376
1376
1389
1389
1410
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1414
1419
1432
1438
1450
1456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      251
272
272
277
277
277
277
277
302
315
315
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1480
1489
1489
1507
1514
1550
1550
1550
1550
1550
1550
1570
                                                                                                                                                                                                           CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                       CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
STRAND
TURN
HELIX
TURN
                                                           ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
                                                                                                                                                                            CARBOHYD
                                           TRANSMEM
                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HELIX
TURN
STRAND
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAND
STRAND
STRAND
TURN
STRAND
HELIX
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAND
HELIX
TURN
STRAND
STRAND
STRAND
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAND
TURN
STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAND
HELIX
TURN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAND
HELIX
HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HELIX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TURN
     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the Buropaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
FUNCTION: NS5A SEEMS TO HAVE A TRANSCRIPTIONAL ACTIVATORY ROLE.
FUNCTION: NS5B IS A RNA-DEPENDENT RNA POLYMERASE THAT PLAYS AN
SESENTIAL ROLE IN THE VIRUS REPLICATION.
CATALYIIC ACTIVITY: Hydrolysis of four peptide bonds in the viral
precursor polyprotein, commonly with Asp or Glu in the P6
position, Cys or Thr in P1 and Ser or Ala in P1'.
CATALYIIC ACTIVITY: N nucleoside triphosphate +
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase;
Core protein; Coat protein; Envelope protein; Helicase; ATP-binding;
Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                SUBDIVIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A LIPOPROTEIN ENVELOPE. THE ENVELOPE CONSISTS OF TWO PROTEINS: E1 AND E2. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.

PTH: THE STRUCTURAL PROTEINS C. E1 AND E2 ARE PRODUCED BY PROTEOLYTIC PROCESSING BY THE HOST SIGNAL PEPTIDASES.

SIMILARITY: THE NS2 PROTEASE BELONGS TO PEPTIDASE FAMILY 039.

SIMILARITY: THE NS3 PROTEASE BELONGS TO PEPTIDASE FAMILY S29.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REMOVED FROM CAPSID PROTEIN C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PROTEASE/HELICASE NS3.
NONSTRUCTURAL PROTEIN NS4A.
NONSTRUCTURAL PROTEIN NS4B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NONSTRUCTURAL PROTEIN NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENVELOPE GLYCOPROTEIN E1 ENVELOPE GLYCOPROTEIN E2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CELLULAR AMINOPEPTIDASE. CAPSID PROTEIN C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROTEIN P7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSFAC; TO4155; ...
InterPro: IPR001410; DEAD.
InterPro: IPR001521; HCV_ccapsid.
InterPro: IPR001521; HCV_ccapsid.
InterPro: IPR001519; HCV_ccapsid.
InterPro: IPR001519; HCV_ms1.
InterPro: IPR001490; HCV_ms2.
InterPro: IPR001490; HCV_ms3.
InterPro: IPR001490; HCV_ms3.
InterPro: IPR001490; HCV_ms4.
InterPro: IPR001650; HCV_ms5a.
InterPro: IPR001994; RNA_pol_pSvir.
Pfam; PF01184; HCV_core; I.
Pfam; PF01184; HCV_ms1; I.
Pfam; PF011860; HCV_ms1; I.
Pfam; PF01006; HCV_ms1; I.
Pfam; PF01006; HCV_ms1; I.
Pfam; PF01006; HCV_ms2; I.
Pfam; PF01016; HCV_ms2; I.
Probm: PF01081; VITAL_RGRP; I.
Probm: P10186062; HCV_ms1; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M67463; AAA45534.1; -. PIR; A36814; GNWVCH. PDB: 1HEI: 25-NOV-98. PDB: 1A1V; 16-FEB-99. PDB: 1A1V; 17-UN-98. MEROPS; S29.001; -. MEROPS; U39.001; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
CHAIN
THILLITITIES AND BURNESS BEAR AND BURNESS BURN
```

```
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entitles requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE—91172826; Pubhed=1848704;

MEDLINE—1848704; Pubhed=1848704;

MEDLINE—1848704;

MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           01-A0G-1992 (Rel. 23, Created)
01-A0G-1992 (Rel. 23, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
16-SEP-2003 (Rel. 42, Last annotation update)
66-60me polyprotein [Contains: Capsid protein C (Core protein) (P22);
67-60 (GP60) (NSI); Protein P7; Nonstructural protein NS2 (P21)
6C 3.4.22.-); Protease/helicase NS3 (P70) (Hepacivirin)
6C 3.4.21.98); Nonstructural protein NS4A (P4); Nonstructural protein NS5A (P56); Nonstructural protein NS5A (P56); CP0) (RNA-directed RNA polymerase) (EC 2.7.7.48)].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (RNA)(N).
SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID COVERED BY A SUBUNIT: THE VIRION OF THIS VIRUS IS A NUCLEOCAPSID OF TWO PROTEINS:
PROTEIN M AND GLYCOPROTEIN E. THE NUCLEOCAPSID IS A COMPLEX OF PROTEIN C AND MRNA.
                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus (isolate 1) (HCV).
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                        Score 44; DB 1; Length 3011; Pred. No. 5.7;
                                                                                                                                                                                                                                                                                               1; Indels
                                                                                                                                                                                   AA; 327142 MW; 772CBB29CCD94753 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 3011 AA
                                                                                                                                                                                                                                                                                               0; Mismatches
                                                                                                                                                                                                                                        89.8%;
90.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; M62321; AAA45676.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001410; DEAD.
                                                                                                                                                                                                                                                                                                                                                                          1678 GCVVIVGRIVL 1688
                                                                                                                                                                                                                                     Query Match 89.8
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                              1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PDB; 1AlV; 16-FEB-99.
PDB; 1HEI; 25-NOV-98.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEROPS; S29.001; ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GNWVC3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NCBI_TaxID=11104;
1598
1606
1614
1622
1627
1635
1640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepacivirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLG_HCV1
P26664;
                                                 TURN
STRAND
STRAND
STRAND
HELIX
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLG_HCV1
ELLLLES
SHIHHH
                                                                                                                                                                                                                                                                                                                                                                                                ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HID DESCRIPTION OF COLUMN TO THE PROPERTY OF THE PROPERTY OF COLUMN TO THE PROPERTY OF THE PRO
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REMOVED FROM CAPSID PROTEIN C BY THE CELLULAR AMINOPEPIDDSE.
CAPSID PROTEIN C (POTEWIIAL).
MAJOR ENVELOPE PROTEIN E (POTEWIIAL).
MAJOR ENVELOPE PROTEIN NSI/EZ (POTEWIIAL).
NONSTRUCTURAL PROTEIN NSI/EZ (POTEWIIAL).
PROTEASE/HELICASE NS3 (POTEWIIAL).
NONSTRUCTURAL PROTEIN NS4 (POTEWIIAL).
NONSTRUCTURAL PROTEIN NS4 (POTEWIIAL).
NONSTRUCTURAL PROTEIN NS4 (POTEWIIAL).
RNA-DIRECTED RNA POLYMERASE (POTEWIIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Polyprotein; Glycoprotein; Transferase; RNA-directed RNA polymerase; Core protein; Coat protein; Envelope protein; Helicase; ATP-binding; Transmembrane; Nonstructural protein; Hydrolase; Serine protease;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
CHARGE RELAY SYSTEM (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 43; DB 1; Length 3011;
Pred. No. 8.7;
1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  65F8C9447FCE5AF9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (GLCNAC. . .)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GLCNAC.
(GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED (GLCNAC, N-LINKED (GLCNAC, N-LINKED (GLCNAC, N-LINKED)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ATP (POTENTIAL).
DECH BOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (
                   InterPro; IPR002166; HCV_NS5a.
InterPro; IPR001656; HCV_RGAP.
InterPro; IPR001650; Helicase_C.
InterPro; IPR007095; RNA_POl_DS_PS.
InterPro; IPR007094; RNA_POl_DS_PS.
Pfam; PF01543; HCV_capsid; l.
Pfam; PF01542; HCV_core; l.
Pfam; PF01560; HCV_core; l.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    N-LINKED
                                                                                                                                                                                                                                                                                                       Pfam; PF01560; HCV_NS1; 1: Pfam; PF01538; HCV_NS2; 1: Pfam; PF01907; HCV_NS3; 1: Pfam; PF01006; HCV_NS4b; 1: Pfam; PF01001; HCV_NS4b; 1: Pfam; PF01506; HCV_NS5a; 1: Pfam; PF00979; Viral, RdRP; 1: ProDom; PD186062; HCV_NS1; 1: ProDom; PD186062; HCV_NS1; 1: ProDom; P0186062; HCV_NS1; 1: Pfam; PF01989; Viral, RdRP; 1: ProDom; P0186062; HCV_NS1; 1: Pfam; PF0186062; 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              87.8%;
81.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               327197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          115
191
383
729
1006
1615
1862
2013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SMART; SM00487; DEXDC;
IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ND-structure.
INIT_MET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ACT_SITE
ACT_SITE
ACT_SITE
ACT_SITE
NP_BIND
SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CARBOHYD
SEQUENCE
```

Transmembrane; Phosphorylation; Magnesium; ATP-binding;

œ

ó

```
Hydrolase;
                                         TRANSMEM
TRANSMEM
                                                                                                                          SEQUENCE
                                                                     TRANSMEM
                                                                                  MOD_RES
METAL
                            TRANSMEM
                                                                                                                                                                                                                                                                                              DPO2_SCHPO
                                                                                                                                                                                                                                                                                                                           ########
                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its modified and properties institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."; Submitted (APPT-2001) to the EMBL/GenBank/DDBJ databases.
-!- CATALYTIC ACTIVITY: AFP + H(2)0 - ADP + phosphate.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: Belongs to the cation transport ATPases family (P-type ATPases). Subfamily IB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN=CDC 1551 / Oshkosh;
Fleischmann R.D., Alland D., Elsen J.A., Carpenter L., White O.,
Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
                                                                                                                                                                                                                                                                                          MEDLINE-9825987; PubMed-9634230;
Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglaneler K., Gas S., Barry C.E. III, Tektia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
"Deciphering the biology of Mycobacterium tuberculosis from the Complete genome sequence.";
Nature 393:537-544(1998).
                                                                                                                         30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
31-SEP-2003 (Rel. 42, Last annotation update)
Probable cation-transporting P-type AlPase C (EC 3.6.3.-).
CTPC OR RV3-270 OR MT3370 OR MTCY71.10.
Wycobacterium tuberculosis, Actinobacteridae, Actinomycetales;
                                                                                                                                                                                                                       Corynebacterineae; Mycobacteriaceae; Mycobacterium
NCBI_TaxID=1773;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Interpro: IPR006404; Heavy_met_ATPase.
Interpro: IPR005834; Hydrolase.
Pfam; PF001122; E1-E2_ATPase; 1.
Pfam; PF00702; Hydrolase; 1.
PRINTS; PR00119; CATATPASE.
TIGRRAMS; TIGR01512; ATPase-IB2_Cd; 1.
TIGRRAMS; TIGR01525; ATPase-IB_hvy; 1.
TIGRRAMS; TIGR01494; ATPase_P-type; 3.
PROSITE; PS00154; ATPASE_E1_E2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro; IPR006416; ATPase-IB_hvy.
InterPro; IPR001757; ATPase_E1-E2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; Z92771; CAB07083.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL, AE007146; AAK47711.1;
PIR; G70978; G70978.
TIGR; MT3370; -.
TubercuList; Rv3270; -.
                            1678 GCVVIVGRVVL 1688
                                                                                                 STANDARD;
1 GSVVIVGRIVL 11
              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                STRAIN-H37Rv
                                                                                                CTPC_MYCTU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Bishai W.
                                                                                                             P96875;
                                                                     RESULT 7
```

```
RATAINSTORY THE ACTION OF THE 
                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILARITY).

SUBUNIT: DNA POLYMERASE ALPHA-PRIMASE IS A FOUR SUBUNIT ENZYME
(SUBUNITS A, B, C AND D), WHICH IS ASSEMBLED THROUGHOUT THE CELL
CYCLE. THE LARGEST SUBUNIT (SUBUNIT A) HAS DNA POLYMERASE
ACTIVITY, THE TWO SUBLILEST SUBUNITS (SUBUNITS C AND D) HAVE DNA
PRIMASE ACTIVITY. SUBUNIT B BINDS TO SUBUNIT A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALPHA/PRIMASE COMPLEX TO THE CELLULAR REPLICATION MACHINERY (BY
                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                             SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                      Score 37; DB 1; Length 718; Pred. No. 30;
                                                                                                                            POTENTIAL.
PHOSPHORYLATION (BY SIMILARI'
MAGNESIUM (BY SIMILARITY).
MAGNESIUM (BY SIMILARITY).
W; BSD6C93AFE636315 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                         1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
Probable DNA polymerase alpha subunit B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     574 AA
                                                                                                                                                                                                                                                                                                                                                                                                                             3; Mismatches
                                                                   POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                  76495 MW;
                                                                                                                                                                                                                                                                                                                                                  75.5%;
                                                                                                                                                                                                                                                                                                                    Ouery Match
Best Local Similarity 63...
7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          286 GSVVVRGRVVV 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 GSVVIVGRIVL 11
                                                                                                                                                             408
610
614
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Schizosaccharomyces.
NCBI_TaxID=4896;
                                                                                                                                                                                                                                            614 (718 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
Complete proteome
                                                            360
616
677
408
610
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DP02_SCHPO
074946;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SPCC553.09C
```

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                        09RPD4;
                                                                                                                                                                                                                                                                                 RESULT 10
MCH_METIR
                                                                                                                                                                                                                                                                                                           DR
DR
DR
DR
DR
DR
DR
DR
DR
                                                                                                                                                                                                                                            q
                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                 ö
                          SIMILARITY).
-!- SIMILARITY: BELONGS TO THE DNA POLYMERASE ALPHA SUBUNIT B FAMILY.
                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSFERRING ELECTRONS FROM SUCCINATE TO UBIQUINONE (COENZYME O)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 subunit, mitochondrial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T., Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S., Agels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A. Rice P., Skelton J., Walsh S., Whitchead S., Barrell B.G.; "The nucleotide sequence of Saccharomyces cerevisiae chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Saccharomyces cerevisiae (Baker's yeast).
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                                                                                                                                                                                                ö
            -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-DEPENDENT MANNER (BY
                                                                                                                                                                                                                                                     Score 35; DB 1; Length 574;
Pred. No. 58;
                                                                                                                                                                                                                                                                                 0; Indels
                                                                                                                                                                                    PIR, 1413-37, 1413-37, 1900, -.

GeneDB_SPOWDe; SPCC553.09c, -.

Pfam; PF04.058; DNA_pol_alpha_B; 1,

DNA replication; Nuclear protein; Phosphorylation,

SEQUENCE 574 AA; 64184 MW; 18C17FCFA7AA098A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-SEP-2003 (Rel. 42, Last annotation update)
Putative succinate dehydrogenase cytochrome B
                                                                                                                                                                                                                                                                                                                                                                                                 196 AA.
                                                                                                                                                                                                                                                                                 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
SUBCELLULAR LOCATION: Nuclear
                                                                                                                                                                EMBL; AL023704; CAA19261.1; -. PIR; T41395; T41395.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; 249702; CAA89756.1; -.
                                                                                                                                                                                                                                                        71.48;
77.88;
                                                                                                                                                                                                                                                                    Best Local Similarity 77.8
Matches 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S54580; S54580.
S0004724; YMR118C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nature 387:90-93(1997).
                                                                                                                                                                                                                                                                                                                                   205 VVVVGRIVV 213
                                                                                                                                                                                                                                                                                                          3 VVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN-S288c / AB972;
PubMed-9169872;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KMR118C OR YM9718.17C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID-4932;
                                                                                                                                                                                                                                                                                                                                                                                                 YM07_YEAST
Q04487;
                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                      YMO7_YEAST
                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE-99412275; PubMed-10482517;
Vorholt J.A., Chistoserdova L.V., Stolyar S.M., Thauer R.K.,
Lidstrom M.E.,
Lidstrom M.E.,
Distribution of tetrahydromethanopterin-dependent enzymes in
methylotrophic bacteria and phylogeny of methenyl
tetrahydromethanopterin cyclohydrolases.";
J. Bacteriol. 181:5750-5757(1999)
-:-FUNCTION: Catalyzes the hydrolysis of methenyl-H(4)MPT(+) to N(5)-
formyl-H(4)MPT (By similarity).
-:-CATALYTIC ACTIVITY: 5,10-methenyl-5,6,7,8-tetrahydromethanopterin
+ H(2)0 - N(5)-formyl-5,6,7,8-tetrahydromethanopterin
-:-PATHWAY: H(4)MPT-dependent pathway of formaldehyde oxidation;
                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bactéria: Proteobactéria: Alphaproteobacteria; Rhizobiales;
Methylocystaceae; Methylosinus.
NCBL_TaxID-426;
                                                                                                                                                                                                                                                                                                                                                             Score 34; DB 1; Length 196;
Pred. No. 34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N(5),N(10)-methenyltetrahydromethanopterin cyclohydrolase (EC 3.5.4.27) (Methenyl-H4MPT cyclohydrolase) (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                     2; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    third step.
-!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
-!- SIMILARITY: BELONGS TO THE MCH FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    260 AA; 27622 MW; BD647C0DBF03C6A8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    260 AA
                                                                                                                                                                                                                                                                                                                                                                                                                   3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HAMAP; MF_00486; -; 1.
InterPro; IPR003209; Cyclohydrolase.
Pf0289; MCH; 1.
Probom; PF011637; Cyclohydrolase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hydrolase; One-carbon metabolism.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; AF162786; AAD56174.1; -.
                                                                                                                                                                                                                                                                                                                                                                  69.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Methylosinus trichosporium
                                                                                                                                                                                                                                                                                                                                                               Query Match 69.4
Best Local Similarity 54.5
Matches 6; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         179 GSIVLAGTLVL 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10LM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MCH_METTR
```

```
ENV_EIAV2
P22428;
                          CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CHAIN
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TRANSMEM
TRANSMEM
TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
             CARBOHYD
                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CARBOHYD
             FT
                                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for connection entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                             ö
                                                                                                                                                                                                                                    01-AUG-1991 (Rel. 19, Created)
01-AUG-1991 (Rel. 19, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein greunsor (Coat polyprotein) [Contains: Coat protein GP90; Coat protein GP45].
                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                   SEGUENCE FROM N.A.
PEDLINE-BOR72070; PubMed-2825406;
PAYDE S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J.,
Montelaro R.C.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "Antigenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes reported for sequential isolates of HIV.";
                                             ;
;
     Score 34: DB 1; Length 260;
Pred. No. 43;
3: Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                 Equine infectious anemia virus (clone P3.2-1) (EIAV).
Viruses, Retroid viruscs; Retroviridae; Lentivirus.
NCBI_TaxID=11666;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-LINKED GGLCNAC...)
                                                                                                                                                                                                       859 AA.
     69.48;
60.08;
Query Match
Best Local Similarity 60.0
Matches 6; Conservative
                                                                                                                                                                                                     STANDARD;
                                                                                                              121 GSVQVVGRVL 130
                                                                            1 GSVVIVGRIV 10
                                                                                                                                                                 RESULT 11

DAY EINV. EIAN1

DO 1-AUG-1991

DT 01-AUG-1991

DT 01-AUG-1991

DT 01-AUG-1991

DT 01-AUG-1991

DE GP96; Codf

GN ENV.

I 16-CT-2001

EQUIDE INF

OX NCBL_TAXID-

RR PRAYE & S.L.

RA MONCELATO.

RE CONTROLLINE SWISS-

CC This SWISS-

CC This SWISS-

CC This SWISS-

CC THIS SWISS-

CC C ARBOHYD

FT CARBOHYD

FT CA
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss institute of Bioinformatics and the EMEL outsitution the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lab-sib.ch).
                                                                                                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-AUG-1991 (Rel. 19, Created)
01-AUG-1991 (Rel. 19, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein preursor (Coat polyprotein) (Contains: Coat protein
GP90; Coat protein GP45).
                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
BEDLINE-BROT2070; PubMed-2825406;
Payne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J.,
Montelaro R.C.;
557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
859 AA; 97140 MW; F4AOC071396DA867 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; M18386; AAA66408.1; -
PIR: B34027; VCLJE2
InterPro: IPR001027; Gp45_EIAV.
InterPro: IPR0010361; Gp50_EIAV.
Pfam: PF001045; EIAV_GP45;
Pfam: PF001045; EIAV_GP45;
Coat profein; Glycoprotein; Polyprotein; Transmembrane; Signal.
SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Antigenic variation and lentivirus persistence: variations in envelope gene sequences during ELAV infection resemble changes reported for sequential isolates of HIV."; virology 161:321-331(1987).
                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                       69.4%; Score 34; DB 1; Length 859; 77.8%; Pred. No. 1.3e+02; ive 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Equine infectious anemia virus (clone P3.2-2) (EIAV). Viruses; Retroid viruses; Retroviridae; Lentivirus. NCBI_TaxID=11667;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ENV POLYPROTEIN,
COAT PROTEIN GP90.
COAT PROTEIN GP45.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL.

N-LINKED (GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         859 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N-LINKED
N-LINKED
                                                                                                                                                                                                 Best Local Similarity 77.8%;
Matches 7: Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          22
8888 25
855 25
861 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87 25
87
                                                                                                                                                                                                                                                                                                                                                                 1 GSVVIVGRI 9
```

```
SEQUENCE FROM N.A.
MEDLINE-87236196; PubMed-3035786;
                                                                                                                                                                                                                                                                                                        XX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /lrology 158:300-312(1987).
      214
233
244
346
346
368
399
406
411
422
483
483
550
557
97066 7
                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity 77.6.
Best Local 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        | |:|||||
| 801 GLVIIVGRI 809
                                                                                                                                                                                                                                                                                                                                                                                                                                1 GSVVIVGRI 9
           proviral DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ENV_EIAV9
P11306;
                            CARBOHYD
CARBOHYD
                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                 CARBOHYD
SEQUENCE
                                                                                           CARBOHYD
                                                                                                                 CARBOHYD
                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                             CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
ENV_EIAV9
             HERE THE TENT OF T
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bloinformatics and the EMBL outstation-the European Bloinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-AUG-1991 (Rel. 19, Created)
01-AUG-1991 (Rel. 19, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein precursor (Coat polyprotein) [Contains: Coat protein GP90; Coat protein GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDIJNE-88072070; PubMed-2825406; Pagnara B.R., Rwambo P., Issel C.J., Pang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J., Panne S.L., Fang F.D., Liu C.P., Dhruva B.R., Rwambo P., Issel C.J., Montelaro R.C., "Antiqenic variation and lentivirus persistence: variations in envelope gene sequences during EIAV infection resemble changes virology 161:321-331(1987).
    (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Polyprotein; Transmembrane; Signal
                                                                                                                                                                                                                                                       Score 34; DB 1; Length 859;
Pred. No. 1.3e-02;
1; Mismatches 1; Indels
                                                                                                                                                                                                 LINKED (GLCNAC. . .) (PC
D86E4E1712E39B32 CRC64;
  N-LINKED (GLCNAC...) (F N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC...) (1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Equine infectious anemia virus (clone P3.2-3) (EIAV).
Viruses: Retroid viruses; Retroviridae; Lentivirus.
NCBL_TaxID=11668;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
ENV POLYPROTEIN.
COAT PROTEIN GP90.
COAT PROTEIN GP45.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M18387; AAA66409.1; ...
PIR; C34027; VCLJE3.
InterPro; IPR001027; Gp45_EIAV.
InterPro; IPR001361; Gp90_EIAV.
Pfam; PF01045; EIAV_GP45; 1.
Coat protein; Glycoprotein; Polyf SIGNAL
                                                                                                                                                                                                                   3
                                                                                                                                                                                                                                                         69.4%;
                                                                                                                                                                                                                   97188
                                                                                                                                                                                                                                                  Query Match 69.4
Best Local Similarity 77.8
Matches 7; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
    3968
3999
3999
3999
4411
4483
4483
757
757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22
8859
8859
8859
8835
8835
1112
1141
1148
1186
                                                                                                                                                                                                                                                                                                                                                                    | |:|||||
| 801 GLVIIVGRI 809
                                                                                                                                                                                                                                                                                                                                           1 GSVVIVGRI 9
368
399
410
411 44
421 44
483
483
490
550
550
557
752
752
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ENV_EIAV3
P22429;
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                          CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                      CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TRANSMEM
TRANSMEM
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TRANSMEM
TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 13
ENV_EIAV3
  FT
FT
FT
FT
FT
FT
SO
                                                                                                                                                                                                                                                                                                                                                                                 a
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for comnercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-JUL-1989 (Rel. 11, Created)
01-FCB-1996 (Rel. 33, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
ENV polyprotein preursor (Coat polyprotein) [Contains: Coat protein
GP90; Coat protein GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC...) (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rawakami T., Sherman L., Dahlberg J., Gazit A., Yaniv A.,
Tronick S.R., Aaronson S.A.;
"Nucleotide sequence analysis of equine infectious anemia virus
                                                                                                                                                                                                                                                                                                                                                                                             69.4%; Score 34; DB 1; Length 859; 77.8%; Pred. No. 1.3e+02; ive 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REVISIONS TO N-TERMINUS.
Tronick S.R.;
Submitted (NOV-1987) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                               982A9F5A1AD8FA4D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Equine infectious anemia virus (clone 1369) (EIAV). Viruses; Retroid viruses; Retroidae; Lentivirus. NCBI_TaxID=11670;
```

```
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Burpoen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-92292130; PubMed-1318398;
MEDLINE-92292130; PubMed-1318398;
Perry S.T., Flaherty M.T., Kelley M.J., Clabough D.L., Tronick S.R., Cogtins L., Whetter L., Lengel C.R., Fuller F.;
"The surface envelope protein gene region of equine infectious anemia virus is not an important determinant of tropism in vitro.";
J. Virol. 66:4085-4097(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PNV_EIAVC STANDARD; PRT; 859 AA. p3241, p32541, p01-0cr-1993 (Rel. 27, Created) p1-0cr-1993 (Rel. 27, Last sequence update) p16-0cr-1001 (Rel. 40, Last annotation update) p16-0cr-1001 precursor (Coat polyprotein precursor (Coat polyprotein) [Contains: Coat protein GP90; Coat protein GP45].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                      (POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
(POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Polyprotein; Transmembrane; Signal.
POTENTIAL.
ENV POLYPROTEIN.
COAT PROTEIN GP90.
COAT PROTEIN GP45.
POTENTIAL.
POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (POTENTIAL)
                                                                                                  POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69.4%; Score 34; DB 1; Length 859; 77.8%; Pred. No. 1.3e+02; ive 1; Mismatches 1; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED (GLCNAC. . .) (P
4B4ED8518CD4F364 CRC64;
                            N-LINKED GGCNAC...)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Equine infectious anemia virus (clone CL22) (EIAV).
Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=31675;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL: M87581; AAA43005.1; --
PIR: C41991; VCLJ22.
InterPro; IPR001027; Gp45_E1AV.
InterPro; IPR0010361; Gp90_E1AV.
Pfam; PF01045; EIAV_GP45; 1.
Coat protein; Glycoprotein; PolygsIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 97113 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | |:|||||
| 801 GLVIIVGRI 809
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 GSVVIVGRI 9
816
40
1112
1148
1148
1186
233
244
233
244
406
406
406
406
550
550
550
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 7; Conserv
                                                                                            CARBOHYD
CAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CHAIN
TRANSMEM
TRANSMEM
TRANSMEM
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CHAIN
      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HE THE THE TARK TO BE THE TARK TO BE
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                         (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                     (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                         (POTENTIAL)
                                                                                        (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-LINKED (GLCNAC. ..) (POTENTIAL N-LINKED (GLCNAC. ..) (POTENTIAL) 23E020E80DF334FA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 859;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 34; DB 1; Length 859
Pred. No. 1.3e+02;
1; Mismatches 1; Indels
                                                                             N-LINKED (GLCNAC,
N-LINKED (GL
                                  POTENTIAL. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Search completed: August 30, 2003, 19:13:51 Job time : 1.54479 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                69.4%;
77.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity 77.00.
7; Conservative
1 GSVVIVGRI 9
                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
CARBOHYD
CARBOHYD
                                                                               CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                            CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
SEQUENCE
                            TRANSMEM
TRANSMEM
                                                                                                                                                                                                                                                                                                                          CARBOHYD
                                                                                                                                                                                                                        CARBOHYD
                                                                                                                                                                                                                                               CARBOHYD
                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
```

ð

```
hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STEATH AND CATEGORY OF PROPERTY OF PROPERT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus type 1b.
Viruses; ssRNA positive strand viruses, no DNA stage: Flaviviridae;
Hepacivirus.
    068227
068227
068227
068207
068216
068221
068221
068220
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
068210
06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bartenschlager R.;
Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
068210
068227
068229
068216
068216
068221
068221
068221
068232
068232
068233
                                                                                                                                                                                                                                                                         068238
068237
068237
                                                                                                                                                                                                                                                                                                                                      068230
081577
081577
081577
081582
081578
081578
081578
081583
081583
081573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAY 2000 (TrEMBLrel, 13, 01-MAY-2000 (TrEMBLrel, 13, 01-MAR-2003 (TrEMBLrel, 23, Genome polyprotein,
  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          [2]
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-31647;
  STRAIN-RB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-RB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 090P06
090P06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 1
09qp06 hepatitis c
09qp61 hepatitis c
081760 hepatitis c
09tfn8 mucosal dis
19914 hepatitis c
1985 hepatitis c
1923 hepatitis c
235 hepatitis c
235 hepatitis c
244 hepatitis c
106 hepatitis c
25 hepatitis c
26 hepatitis c
27 hepatitis c
27 hepatitis c
106 hepatitis c
27 hepatitis c
27 hepatitis c
107 hepatitis c
108 hepatitis c
109 hepatitis c
109 hepatitis c
100 hepatitis c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          hepatitis
hepatitis
                                                                                                                                    August 30, 2003, 19:00:22; Search time 2.09905 Seconds (Without alignments) 1352.314 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred. No. 1s the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     099661 b
0981760 b
0981760 b
099168 m
039914 b
039895 b
068209 b
068235 b
068218 b
068244 b
068244 b
                                                                                                                                                                                                                                                                                                                                                                                                                           830525
                 GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                    Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                              830525 seqs, 258052604 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUMMARIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                       using sw model
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SPTREMBL_23:*

1: sp_archea:*
2: sp_bacteria:*
3: sp_fung1:*
4: sp_human:*
5: sp_invertobrate:*
5: sp_amcal:*
5: sp_amcal:*
5: sp_amcal:*
5: sp_phage:*
5: sp_phage:*
5: sp_phage:*
5: sp_phage:*
5: sp_organelle:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               090P06
090P61
081760
090TE5
039914
039895
068209
068235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    068244
068206
068242
068240
068240
                                                                                                                                                                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sp_vertebrate:*
sp_unclassified:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sp_bacteriap:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        sp_rvirus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-26
49
                                                                                                                                                                                                                                                                       1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  sp_archeap:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sp_virus:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3010
3010
3010
3010
4040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                     - protein search,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
100000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
10000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score
                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                              Searched:
                                                                                                                                                                                                                                                                           Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Database
                                                                                                                                             Run on:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Result
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
```

```
IPR001490;
                                                                                                               InterPro; InterP
                                                    InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    081760
081760;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             datches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 3
Q81760
                                                    ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ó
InterPro; IPR002866; HCV_NS5a.

DR InterPro; IPR002166; HCV_NS5a.

DR InterPro; IPR007095; RNA_pol_DS_PS.

DR Pfam; PP01542; HCV_core; 1.

DR Pfam; PP01543; HCV_core; 1.

DR Pfam; PP01538; HCV_NS2; 1.

DR Pfam; PF01589; HCV_NS2; 1.

DR Pfam; PF01506; HCV_NS2; 1.

DR Pfam; PF01506; HCV_NS2; 1.

DR Pfam; PF01006; HCV_NS2; 1.

DR Pfam; PF01006; HCV_NS3; 1.

DR Pfam; PF01006; HCV_NS3; 1.

DR Pfam; PF001506; HCV_NS3; 1.

DR Pfam; PF001506; HCV_NS3; 1.

DR Pfam; PF0098; Viral_RdRP; 1.

DR Probom; PD186662; HCV_NS3; 1.

DR PROSTIE; PS50507; RDRP_POSITIVE; 1.

DR PROSTIE; PS50507; RDRP_POSITIVE; 1.

DR PROSTIE; PS50501; RDRP_POSITIVE; 1.

DR POLYPROTEIN; RNA-directed RNA POLYBREASE; TransMembrane.

FT CHAIN 192 383 GLYCOPROTEIN E1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRAIN-274933RU:
CSTRAIN-274933RU:
CSTRAIN-274933RU:
CSTRAIN-274933RU:
CSTRAIN-274933RU:
CSTRAIN-274933RU:
CSTRAIN-274933RU:
CSTRAIN-274938RU:
CSTRAIN-27493BU:
CSTRAIN-2749BU:
CSTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 01-MAY-2000 (TYEMBLEEL 13, Created)
01-MAY-2000 (TYEMBLEEL 13, Last sequence update)
01-MAR-2003 (TYEMBLEEL 23, Last annotation update)
Genome polyprotein.
Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              183 GLYCOPROTEIN E1.
746 GLYCOPROTEIN E2.
809 P7 PEPTIDE.
1026 NS2 PROTEINASE.
1157 NS3 PROTEINASE.
1157 NS3 PROTEINASE.
1157 NS4 PROTEINASE.
1972 NS4 PROTEIN.
1972 NS4 PROTEIN.
1973 NS5 PROTEIN.
1973 NS5 PROTEIN.
1973 NS5 PROTEIN.
1974 NS5 PROTEIN.
1975 NS5 PROTEIN.
1976 NS5 PROTEIN.
1977 NS5 PROTEIN.
1976 NS5 PROTE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro; IPR001410; DEAD.
InterPro; IPR00252; HCV_capsid.
InterPro; IPR002521; HCV_core.
InterPro; IPR002519; HCV_env.
InterPro; IPR002519; HCV_NSI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                810
1027
1658
1712
1712
1973
2419
3010 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID+11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 090P61
090P61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CHAIN
CHAIN
CHAIN
CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
```

```
A STADIAL AND STADIAL STATE ST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE: PS50507; RDRP_POSITIVE: 1.
PROSITE: PS50521; RDRP_VIRAL; 1.
ATP-binding: Coat protein; Envelope protein; Glycoprotein; Helicase; Hydrolase; Nonstructural protein; Polyprotein;
RNA directed RNA polymearase; Transferase; Transmembrane.
SEQUENCE 3010 AA; 327068 MW; 9105F69483DD5BBA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6; tive 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT; 3010 A.A.
                                                                         InterPro: IPR001400; HCV_NG4b.
InterPro: IPR001400; HCV_NG4b.
InterPro: IPR001266; HCV_RG5b.
InterPro: IPR001266; HCV_RG5b.
InterPro: IPR001060; HCV_RG4b.
InterPro: IPR001060; HM_Dol_DS_PS.
InterPro: IPR001094; RNA_Dol_DS_PS.
InterPro: IPR001094; RNA_Dol_DS_PS.
InterPro: IPR001094; RNA_Dol_DS_PS.
InterPro: IPR001094; RNA_Dol_DS_PS.
IPR001590; HCV_COTE; 1.
IPR001590; HCV_COTE; 1.
IPR001590; HCV_NS2; 1.
IPR001590; HCV_NS2; 1.
IPR001590; HCV_NS2; 1.
IPR001590; HCV_NS3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV_capsid.
HCV_core.
HCV_env.
HCV_NS1.
HCV_NS2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV_NS4a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, 01-NOV-1996 (TrEMBLrel. 01, 01-MAR-2003 (TrEMBLrel. 23,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1678 GSVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local Similarity 100.
hes 11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 GSVVIVGRIVL 11
IPR004109;
IPR000745;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                IPR002519;
IPR002531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=11103;
```

ö

us-09-965-594-26.rspt

```
Lai V.C., Zhong W., Skelton A., Ingravallo P., Vassilev V.,
Donis R.O., Hong Z., Lau J.Y.;
"Generation and characterization of a hepatitis C virus NS3 protease-
dependent bowine viral diarrhea virus.";
J. Virol. 74:6339-6347(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                Mucosal disease virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Virol. 74:6339-6347(2000).
[2] SEQUENCE FROM N.A.
Lal V.C.H., Hong Z.;
Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
EMBL, AF266278; AaF82566.1;
HSSP; P26663; LJXP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-077-2000 (TrEMBLrel. 15, Created)
01-077-2000 (TrEMBLrel. 15, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 4040 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INTERPRO; 191.001.
INTERPRO; IPR000280; CDVir_endptseP80.
INTERPRO; IPR0001409; HCV_NS3.
INTERPRO; IPR0014109; HCV_NS3.
INTERPRO; IPR001505; HCV_NS3.
INTERPRO; IPR001505; HG1CASE_C.
INTERPRO; IPR001005; Myb_DNA_binding.
INTERPRO; IPR001005; Myb_DNA_binding.
INTERPRO; IPR0010958; RNA_pOl_DS_PS.
INTERPRO; IPR007094; RNA_pOl_DS_PS.
INTERPRO; IPR007094; RNA_pOl_PSVIR.
PÉAM; PF00271; HCV_LS_RNA_POL_PSVIR.
PÉAM; PF00271; HCAL_RGRP; I.
PÉAM; PR00721; HCAL_RGRP; I.
PRINTS; PR00749; CDVENDOPTASE.
SMART; SM00489; HELICC; I.
SMART; SM00489; HELICC; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       [1]
SEQUENCE FROM N.A.
MEDLINE-20323484; PubMed-10864644;
MEDLINE-20323484; PubMed-10864644;
                                      pfam; pr01542; HCV_core; 1. pfam; pr01542; HCV_core; 1. pfam; pr01560; HCV_NS1; 1. pfam; pr01550; HCV_NS2; 1. pfam; pr02007; HCV_NS3; 1. pfam; pr01006; HCV_NS4; 1. pfam; pr01006; HCV_NS4; 1. pfam; pr01506; HCV_NS5; 1. pfam; pr01506; HCV_NS5; 1. pfam; pr01506; HCV_NS5; 1. pfam; pr001504; helicase_C; 1. pfam; pr001504; VIral_RGRP; 1. propro. pr01505; HCV_NS1; 1. pr0150; Pr01505; HCV_NS1; 1. pr01505; HCV_NS2; 1. pr01505; HCV_NS2; 1. pr01505; HCV_NS2; 1. pr01505; HCV_NS2; 1. pr01505; H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1678 GSVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genome polyprotein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11099;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MCROPS; S31.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q91FH8
Q91FH8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
091FH8
               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DORAN BERTARAN BERTAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0
                                                              InterPro; IFR007094; RNA_pol_DS_PS.

R InterPro; IFR007094; RNA_pol_DS_PS.

R InterPro; IFR007094; RNA_pol_DS_PS.

R InterPro; IFR007094; RNA_pol_DS_PS.

R Pfam; PF01543; HCV_cars; 1.

R Pfam; PF01549; HCV_cars; 1.

R Pfam; PF01560; HCV_NS1; 1.

R Pfam; PF01000; HCV_NS3; 1.

R Pfam; PF01000; HCV_NS3; 1.

R Pfam; PF01000; HCV_NS5; 1.

R Pfam; RD1806CS; HCV_MS5; 1.

R PROSITE; PS50507; RDRP_POSITIVE; 1.

R PROSITE; PS50507; RDRP_POSITIVE; 1.

R PROSITE; PS50507; RDRP_VIRAL; 1.

R PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; saRNA positive strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,
Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 49; DB 12; Length 3010; 100.0%; Pred. No. 6.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT; 3010 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA_pol_DS_PS.
RNA_pol_PSvir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HCV_RdRP.
Helicase_C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR002522; HCV_capsid.
InterPro; IPR002521; HCV_core.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR002519; HCV_env.
InterPro; IPR002531; HCV_NS1.
InterPro; IPR002518; HCV_NS2.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR004109; HCV_NS3.
InterPro; IPR002868; HCV_NS5a.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1678 GSVVIVGRIVL 1688
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Local Similarity 100.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IPR000745;
IPR001490;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               IPR007094;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Genome polyprotein.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Repactvirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-HCVT145;
                                              InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q9DTE5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4

OSDIES

OSDI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
   g
```

ö

Gaps

```
J. Gen. Virol. 76:211-215(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                         Local Similarity 90.9 ses 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             20 GSVVIVGRIIL 30
                                                                                                                                                                                                                                                     1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                          8 GSVVIVGRIIL 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Fragment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                     O9PXP5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                068209
                                                                                                                                                                                                                                                                                                                                                                                                  Q9PXP5
                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                         RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 9
                                                                                                                                                                                                                                                                                                                                                                              09PXP5
                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AC CON NO CON NO
          SOFFER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            셤
                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ö
                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Sequence analysis of hepatitis C virus variants producing discrepant results with two different genotyping assays.";
J. Med. Virol. 53:237-244(1997).
EMBL: AF007519; AB62970.2;
InterPro; IPR00745; HCV.MS4a.
Pfam: PF01006; HCV_NS4a: 1.
PROSITE; PS50507; RDRP_POSITIVE; 1.
PROSITE; PS50521; RDRP_VIRAL; 1.
PROSITE; PS00531; RNASE_12_2; 1.
ATP-binding; Helicase; Hydrolase; Nonstructural protein; Polyprotein; RNA-directed RNA polymerase; Transferase.
SEQUENCE 4040 AA; 453073 MW; ADE87791D055B9DC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Sequence analysis of Hepatitis C virus variants producing discrepant results with two different genotyping assays.";
                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Prescott L.E., Berger A., Pawlotsky J.-M., Conjeevaram P., Pike I.,
Simmonds P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN-FD_1;
BDLINE-198032593; PubMed-9365889;
Prescott L.E., Berger A., Pawlotsky J.M., Conjeevaram P., Pike I.,
Simmonds P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .;
0
                                                                                                                                                                                         ö
                                                                                                                                          Query Match
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 11; Conservative 0; Mismatches 0; Indels 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 48; DB 12; Length 88;
Pred. No. 0.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       88 88
88 AA; 9750 MW; BF7B5198B317B6E0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                 Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                              88 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         89 A.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                               01-JNN-1998 (TrEMBLrel. 05, Created)
01-JNN-2002 (TrEMBLrel. 21, Last Sequent
01-OCT-2002 (TrEMBLrel. 22, Last annotat
Non-structural protein 4a/b (Fragment).
NS4A/B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-JAN-1998 (TrEMBLrel. 05, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequen
01-MAR-2002 (TrEMBLrel. 20, Last annota
Non-structural protein 4a/b (Fragment).
                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98.08;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity 90.5
                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                  1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                        10 GSVVIVGRIVE 20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVI, 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                6 GSWIVGRIIL 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    039895
                                                                                                                                                                                                                                                                                                                                                                          039914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NS4A/B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                 RESULT 6
039914
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
```

```
;
                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
MEDLINE-94245087: PubMed-7514558;
Tanaka T., Tsukiyama-Kohara K., Yamaguchi K., Yagi S., Tanaka S.,
Hasegawa A., Ohta Y., Hattori N., Kohara M.;
"Significance of specific antibody assay for genotyping of hepatitis C
                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Viruses; sRNA positive-strand viruses, no DNA stage: Flaviviridae;
Hepacivirus.
NCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MEDLINE=95146953; PubMed=7844535;
Greene W.K., Cheong M.K., Ng V., Yap K.W.;
"Prevalence of hepatitis C virus sequence variants in South-East
Asia.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                             ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2002 (TrEMBLrel. 11, Last annotation update)
NON-structural protein NS4-GROUP II HCV-specific antigen C14-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 102;
                                                                                                                                                                                                     Score 48; DB 12; Length 89;
Pred. No. 0.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0; Indels
                                                                                                                                                                                                                                                           0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE 102 AA; 11419 MW; 08124C19CF367F06 CRC64;
J. Med. V4rol. 0:0-0(1997).

EMEL; AF007500; AAB62951.1; -
Interp.: IPRO0145; HCV_NS4a.

Pfam; PF01006; HCV_NS4a.

1 1.

NON_TER 1 1.

NON_TER 89 89

SEQUENCE 89 AA; 9748 MW; 581BBBC8A3EA8B5C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DC-2001 (TEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98.0%; Score 48; DB 12;
90.9%; Pred. No. 0.35;
Live 1; Mismatches 0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      138 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           102 AA
                                                                                                                                                                                                                                                     1; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatology 19:1347-1353(1994).
InterPro; IPR000745; HCV_NS4a.
Pfam; PF01006; HCV_NS4a; 1.
NON_TER 102 102
                                                                                                                                                                                                       98.0%;
90.9%;
```

Ŋ

```
"Prevalence of hepatitis C virus sequence variants in South-East
                                                                                                                                                                                                                                                                                                                                                                                                                                          138 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             138 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                    EMBL, Uirol, 76:211-215(1995).
EMBL, U14254; AAC53943.1;
HSSP; P27958; IHEI.
Interpro; IPR000745; HCV_NS4a.
Pfam; PF01006; HCV_NS4a; I.
NON_TER 138 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Asia.";
J. Gen. Virol. 76:211-215(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 90.9
nes 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRELIMINARY;
                                                                                                                                                                                                                                                                                                      1 GSWIVGRIVL 11
                                                                                                                                                                                                                                                                                                                             1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBL_TaxID=11103;
                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                              068244;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             068206;
                                                                                                                                                                                                                                                                                                                                                                                                                                          068244
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             968206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rest Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0068206
110 006
007 001
007 001
007 001
007 001
008 NO
                                                                                                                                                                                                                                                                                                                                                                                                                       068244
                                                                                                                                                                                                                                                                                                                                                                                                                                              SO OC BREEFER SO OC CONTRACTOR SO OCCUR SO OCCU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
      õ
                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Greene W.K., Cheong M.K., Ng V., Yap K.W.;
"Prevalence of hepatitis C virus sequence variants in South-East
Asia.";
                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                         98.0%; Score 48; DB 12; Length 138; 90.9%; Pred. No. 0.47; ative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 98.0%; Score 48; DB 12; Length 138; Best Local Similarity 90.9%; Pred. No. 0.47; Matches 10; Conservative 1; Mismatches 0; Indels
                                                                             NON_TER 1 1 1 NON_TER 138 138 SEQUENCE 138 AA; 15149 WW; DBAE62A0FE9E2D57 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Interpro; IPR000745; HCV_NS4a.
Pfam: PF01006; HCV_NS4a: 1.
NON_TER 1 1.
NON_TER 138
SEQUENCE 138 Aa: 15104 MW; 585DC5A627D0F3E3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                               01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TIEMBLrel. 01, Created)
1-NOV-1996 (TIEMBLrel. 01, Last sequence update)
01-DEC-2001 (TIEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                              138 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               138 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-1D;
MEDLINE-95146953; PubMed-7844535;
Greene W.K., Cheong M.K., Ng V., Yap K.W.;
                                                                                                                                                                                                                                                                                                                                                                                PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN-1b;
MEDLINE-95146953; PubMed-7844535;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                J. Gen. Virol. 76:211-215(1995).
EMBL: U14271; AAC53960.1; -.
HSSP; P27958; 1HEI.
EMBL; U14245; AACS3934.1; -.
HSSP; P27958; 1HEI.
InterPro; IPR000745; HCV_NS4a.
Pfam; PF01006; HCV_NS4a; I.
                                                                                                                                                       Query Match
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                         1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                1 GSVVIVGRIVE 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                52 GSVVIVGRIIL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WCBI_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q68218
Q68218;
                                                                                                                                                                                                                                                                                                                                                                              068235
                                                                                                                                                                                                                                                                                                                                                                                                 068235
                                                                                                                                                                                                                                                                                                                                         RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
                                                                                                                                                                                                                                                                                                                                                                268235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
```

```
.
0
                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
NCBL_TaxID=11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-lb;
MEDLINE-95146953; PubMed-7844535;
Greene W.K., Cheong M.K., Ng V., Yap K.W.;
"Prevalence of hepatitis C virus sequence variants in South-East
                                                                                                                                               ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   98.0%; Score 48; DB 12; Length 138; 90.9%; Pred. No. 0.47; tive 1; Mismatches 0; Indels
                                                                                 Query Match 98.0%; Score 48; DB 12; Length 138; Best Local Similarity 90.9%; Pred. No. 0.47; Matches 10; Conservative 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; U14280; AAC53969.1; -...
HSSP; P27958; IHEI.
InterPro; IPR000745; HCV_NS4a.
Pfam; PF01006; HCV_NS4a; 1.
NON_TER 138 138
SEQUENCE 138 AA; 15118 MM; B7F7EB2733770408 CRC64;
NON_TER 138 138
SEQUENCE 138 AA; 15189 MW; DB78E92DDC67040F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TYEMBLrel. 01, Created)
01-NOV-1996 (TYEMBLrel. 01, Last sequence update)
01-DEC-2001 (TYEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
```

```
NCBI_TaxID=11103;
  ŝ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis C virus.
Viruses: ssRna positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus.
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
       MEDLINE-10;
MEDLINE-10;
Green W.K., Cheong M.K., Ng V., Yap K.W.;
Green W.K., Cheong M.K., Ng V., Yap K.W.;
Prevalence of hepatitis C virus sequence variants in South-East Asia., J. Gen. Virol, 76:211-215(1995).
EMBL; 014422, AAC53931.1; -.
EMBL; 014422, AAC53931.1; -.
EMBL; 014425, AAC53931.1; -.
EMBL; 016425, AAC53931.1; -.
EMBL; 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Greene W.K., Cheong M.K., Ng V., Yap K.W.; 
 "Prevalence of hepatitis C virus sequence variants in South-East
                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98.0%; Score 48; DB 12; Length 138; 90.9%; Pred. No. 0.47; tive 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                           98.0%; Score 48; DB 12; Length 138; 90.9%; Pred. No. 0.47; Live 1; Mismatches 0; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE 138 AA; 15281 MW; CD5B5B3834C6070D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                068242:
01-NOV-1996 (TrEMBLrel. 01, Created)
10-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DBC-2001 (TrEMBLrel. 19, Last annotation update)
Nonstructural protein (Fragment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-NOV-1996 (TIEMBLrel. 01, Created)
01-NOV-1996 (TIEMBLrel. 01, Last sequence update)
01-DEC-2001 (TIEMBLrel. 19, Last annotation update)
Nonstructural protein (Frägment).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     138 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               138 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-1b;
MEDLINE-95146953; Pubmed-7844535;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    J. Gen. Virol. 76:211-215(1995).
EMBL; U14278; AAC53967.1; -.
HSSP; P27958; 1HEI.
Interpro; IPR000745; HCV_NS4a.
Pfam; PF01006; HCV_NS4a; 1.
                                                                                                                                                                                                                                                                                                                                Query Match 98.0
Best Local Similarity 90.9
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity 90.9
les 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      52 GSVVIVGRIIL 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                                                            1 GSVVIVGRIVI, 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                        138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID-11103;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q68240
Q68240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               068242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 15
068240
1D 06824
AC 06824
DT 01-NO
DT 01-NO
DT 01-DE
DE NONST
GN NS4.
CS HEPAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
Q68242
õ
```

```
RP SEQUENCE FROM N.A.

RC STRAIN-1b;

RX MEDLINE-91146933: Pubmed-7844535;

RA Greene W.K., Cheong M.K., Ng V., Yap K.W.;

RT "Prevalence of hepatitis C virus sequence variants in South-East

RT "Prevalence of hepatitis C virus sequence variants in South-East

RT "Asia.";

R J Gene Virol. 76:211-215(1995).

DR HSSP: P27598: 1HEI.

DR HSSP: P27698: 1HEI.

DR HSSP: P27698: 1HEI.

DR HSSP: P7600745: HCV_NS4a: 1.

FT NON_TER 138

COURTY MACH 138

COURTY MACH 138

A. 15115 MW, 6A042677B354CA7A CRC64;

COURTY MACH 138

A. 15115 MW, 6A042677B354CA7A CRC64;

COURTY MACH 138

A. 15115 MW, 15115 MW, 6A042677B354CA7A CRC64;

COURTY MACH 138

D. SEQUENCE 138 A. 15115 MW, 18804 MACH 138;

Bast Local Similarity 90.9%: Pred. No. 0.47;

MATCHES 10; Conservative 1; Mismatches 0; Indels 0; Gaps

Oy 1 GSVVIVGRIUL 62

Search completed: August 30, 2003, 19:18:23

Job time: 4.09905 seess
```

ö

```
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                                                                                                                                                     SUMMARIES
                                                                                                                                                                                                                              AR207294
HCJ238800
AB049092
                                                                                                                                                                                                                                                       HPCCGENOM
HPCRNA
                                                                                                                                                                                                                                                                                                                                                                                                                                  HCU14241
HCU14242
HCU14243
HCU14244
HCU14246
HCU14246
HCU14246
                                                                                                                                                                                                                                                                       AF176573
AR179057
AF268278
                                                                                                                                                                                                                                                                                                                                                                                         E09291
AF007500
                                                                                                                                                                                                                                                                                                                                                                         AF007519
                                                                                                                                                                                                                                                                                               ARI45217
ARI45217
ARI45197
ARI45213
ARI45247
ARI45249
AX481515
ARO37527
                                                                                                                                                                                             E06198
E06394
A91965
AR031992
                                                                                                                                                                                                                                                                                                                                                                                 AR037528
                                                                                                                                                                                                                                                                                                                                                                  E09290
                                                                                                                                                                                                                                                                                                                                                                                                           E11063
                                                                                            em_htgo_mus:*
em_htgo_other:*
                em_htg_inv:*
em_htg_other:*
                                                                                   em_htgo_hum: *
        em_htg_hum:*
                                 em_htg_mus: *
em_htg_pln: *
                                                  em_htg_rod:*
                                                          em_htg_mam:*
em_htg_vrt:*
                                                                                                                                                                              DB
                                                                                                                                                                     Query
Match Length
                                                                                                                                                                                                                                        9033
                                                                                                                                                                                                             100.0
Score
                                                                                                                                                                                              RESULT 1
                                                                                                                                                                                                                                                                                                Result
                                                                                                                                                                                                                                                                             -WODEL-frame+_p2n.model -DEV-xlp
-Q-Cqqq2_1_XCBFQ_spool_VG09965594_frunat_29082003_151919_28110/app_query.fasta_1.2872
-Q-Cqqq21_XCBFQ_spool_VG09965594_frunat_29082003_151919_28110/app_query.fasta_1.2872
-Q-Cqqq2_1_XCBFQ_spool_VG09965594_rep -NMRATCH-0.1 -LOOPCL-0 -LOOPEXT-0
-UNITS-bits -START-1 -END--1 -MATRIX-blosum62 -TRANS-human40.cd1 -LIST-45
-DOCALIGN-200 -THR_SCORE-FCT -THR_MXX-100 -THR_MXN-10 -ALIGN-15 -MODE-LOCAL
-OUTFHT-ptc -NORM-ext -HEAPSIZE-500 -MINIT -DSPBLOCK-100 -LONGLOG
-USER-US09965594_eCGN_1 1_14686_erunat_29082003_151919_28310 -NCPU-6 -ICPU-3
-NO_MARP -LARGEQUERY -NEGCORES-0 -WAIT -DSPBLOCK-100 -LONGLOG
-NGARD -LARGEQUERY -NEGROUS -NARN THEOUT-30 -THREADS-1 -XGAPDP-10 -XGAPEXT-0.5 -FGAPOP-6
-FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELDP-6 -DELEXT-7
                                                        August 30, 2003, 19:18:33; Search time 142.976 Seconds (without alignments) 3147.423 Million cell updates/sec
       GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                        nucleic search, using frame_plus_p2n model
                                                                                                                                                                            2888711 segs, 20454813386 residues
                                                                                                                                                                                            hits satisfying chosen parameters:
                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                  Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                                                                                                                             Minimum DB seq length: 0
Maximum DB seq length: 2000000000
                                                                                         US-09-965-594-26
                                                                                                          1 GSVVIVGRIVL 11
                                                                                                                                                                                                                                                                                                                                                                                9b_ba:*
9b_htg:*
9b_in:*
9b_on:*
9b_pat:*
9b_pat:*
9b_pr:*
9b_pr:*
9b_rr:*
9b_rr:*
9b_vr:*
9b_vr:*
9b_vr:*
9b_vr:*
9b_vr:*
9b_vr:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_pl:*
em_ro:*
em_sts:*
em_un:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           em_in:*
em_mu:*
em_om:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      em_or:*
em_ov:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     em_pat:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               em_ph:*
                                                                                                                                                                                                                                                                                                                                                                         GenEmbl: *
                                                                                                                            BLOSUM62
                                                                                                                                                                                                                                                                        Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           oţ
                                                                                                  Perfect score:
                                                                                                                            Scoring table:
                                         OM protein -
                                                                                                                                                                                             Total number
                                                                                                                                                                                                                                                                                                                                                                         Database :
                                                                                                           Sequence:
                                                                                                                                                                             Searched:
                                                         Run on:
                                                                                           ritle:
```

AR031992 Sequence AR207294 Sequence AJ238800 Hepatitis AD049092 Hepatitis L02836 Hepatitis C D10934 Hepatitis C AR176573 Hepatitis AR176057 Sequence

E06198 cDNA encodi

Description

AF268278 Pestiviru

AR145217 Sequence AR145221 Sequence AR145197 Sequence AR145247 Sequence AR145249 Sequence AR145249 Sequence AR037527 Sequence E09290 DNA encodin AF007519 Hepatitis

E09291 DNA encodin AF007500 Hepatitis E11063 DNA sequenc E11062 DNA sequenc E03389 DNA encodin

014241 Hepatitis Cold 1424 Hepatitis Cold 1425 Hepatitis Cold 1425 Hepatitis Cold 1425 Hepatitis Cold 1425 Hepatitis Cold 1426 Hepatitis Cold 14266 Hepatitis Cold 1426 Hepatitis Cold 142

```
Hepacivitus.

S Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and Hayashi,N. Era. Corp. BY THE SAME AND ITS PRODUCTION PROM PEPATITIS C VIRUS, POLYPEPTIDE CODED BY THE SAME AND ITS PRODUCTION PAINTISHI KASEI CORP.

MITCHISTAIN KASEI CORP.

OS (hepatitis C virus)
PN JP 1994000086-A/38
PP 10-3A-1994
PF 07-CCT-1994
PF 07-CCT-1995
PF 07-CCT-1994
PF 07-CCT-19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     circular PAT 22-JAN-2000
             DV0337.1 2.1.2...
JP 199400086-A/38.
Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               C12N15/85,C12P21/O2//A61K39/29,(C12P21/O2,C12R1:91),C07K99:00;
strandedness: Double;
topology: Linear;
ant1-sense: No;

    .608
/product-'the peptides reacting specifically and

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         immunochemically with the serum of hepatitis
                                                                                                                                                                                                                                                                                                                                                                                                                                                  TERANISHI YUTAKA, HAYASHI NORIO
C12N15/51,C07K7/06,C07K7/08,C07K7/10,C07K13/00,C12N5/10,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 618
111
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
1. .618
//organ.sm="Hepatitis C virus"
//mol_type="genomic RNA"
//db_xref="taxon:11103"
a 184 c 179 g 120 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 (bases I to 7475)
Attwood, M.R. and Hurst, D.N.
MITVIRAL PEPTIDE DERIVATIVES
PATENT: WO 9822496-A I 28-MAY-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A91965 7475 bp
Sequence 1 from Patent W09822496.
A91965
A91965.1 G1:6740811
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tient'.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-26 (1-11) x E06394 (1-618)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    type C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source: clone-N13-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0.238
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               unidentified unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              unidentified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                  PI TERAN
C12N15/11,
PC C12N15/11,
PC C12N16
CC Stran
CC antl-c
CC antl-c
FH Key
FH Key
FT mat_P
FT FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ..
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BASE COUNT
                                                                  ORGANISM
                                                                                                                               REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
A91965
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURES
                                                                                                                                                                                             TITLE
                                                                                                                                                                                                                                                                             COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PAT 29-SEP-1997
                                                                                                                                                                         Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                        PAT 29-SEP-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                C12R1:19), (C12N5/10,C12R1:91), (C12P21/02,C12R1:19), (C12P21/02,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PI TERANISHI YUTAKA, HAYASHI NORIO
C C12N15/51,007K7/06,C07K7/08,C07K7/10,C07K13/00,C07K15/12, PC
C12N1/21,C12N5/10,
PC
C12N15/11,C12N15/70,C12N15/85,C12P21/02//A61K39/00,A61K39/29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                  172794, PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9. .608
/product-'the peptides reacting specifically
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       immunochemically with the serum of hepatitis
                                                                                                                                                                                                                                                         Hayashi,N.
GENE OR DNA FRAGMENT DERIVED FROM HEPATITIS C VIRUS, POLYPEPTIDE
CODED BY THE SAME AND ITS PRODUCTION
PATENT: JP 1994000085-A 38 11-JAN-1994;
MITSUBISHI KASEI CORP
                                                                                                                                                                                                               1 (bases 1 to 618)
Seki,M., Honda,Y., Takahashi,K., Murakami,T., Teranishi,Y. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      332329, PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PI SEKI MANOTO, HONDA YOSHIKAZU, TAKAHASHI KAZUNOBU, PI
MURAKAMI TOMOKO,
                                                                                                                                                                                                                                                                                                                        CDNA encoding genes derived from hepatitis C virus.
E06394
                      18 bp RNA linear encoding genes derived from hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        618
00
00
00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.-618
/organism-"Hepatitis C virus"
/mol_type-"genomic RNN"
/db_xref-"taxon:11103"
a 184 c 179 g 120 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           type C
    patient'.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-26 (1-11) x E06198 (1-618)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              strandedness: Double;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            *source: clone=N13-1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    topology: Linear;
anti-sense: No;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.238
49.00
100.008
100.008
                                                                                      E06198.1 GI:2174385
                                                                                                           JP 1994000085-A/38
                                                                                                                               Hepatitis C virus
Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (C12N1/21,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          135
                                                 CDNA en
E06198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                      LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ..
Q
                                                                                                                                                                                                               REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                        JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
                                                                                                                                                                                                                                                                             TITLE
                                                                                                                                                                                                                                                                                                                                                                  COMMENT
  E06198
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8
```

m

```
Location/Qualifiers
1. .7475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 2 (bases 1 to 9033)
Bartenschlager, R.
Direct Submission
                                                                              3.53
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1. ,9033
                                                                                                    Percent Similarity:
Best Local Similarity:
                                    1853 a
                                                                                                                                                                                                                                                                                                                                                                                                                                            99370154
                                                                                                                                                                                                                                                                                                                                                                                                                                                      10438800
                                                                     Alignment Scores:
                                                                                                                          Query Match:
DB:
                                                                                                                                                                                                                                                              DEFINITION
ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                         SOURCE
ORGANISM
                                   BASE COUNT
ORIGIN
                                                                               Pred. No.:
                                                                                                                                                                                                                              RESULT 6
HCJ238800
                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
MEDLINE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                        PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURES
   FEATURES
                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                    PAT 29-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAT 20-JUN-2002
                                                                                                                                                                                                                                                                                                                                                                       1 (bases 1 to 7475)
Attwood, M. Richard., Hurst, D. Nigel., Jones, P. Stephen.,
Kay, P. Brittain., Rayhham. T. Michhael. and Wilson, F. Xavier.
Peptidyl inbitors of viral proteases
Patent: US 5866684-A 1 02-FEB-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 (bases 1 to 7475)
Attwood, M.Richard., Hurst, D.Nigel., Jones, P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.
Antiviral medicaments
Patent: US 6372883-A 1 16-APR-2002;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                              7475
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7475
111
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                    Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                   1 GlySerValValIleValGlyArgIleValLeu 11
                               /organism="unidentified"
/mol_typc="genomic DNA"
/db_xref="taxon:32644"
i 1916 c 2039 g 1667 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ب
                                                                                                                Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1667
                                                                                                                                                            Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 1 from patent US 6372883.
                                                                                                                                                                                                                                                                                Sequence 1 from patent US 5866684. AR031992. G:5946281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-965-594-26 (1-11) x AR031992 (1-7475)
                                                                                                                                                                                            US-09-965-594-26 (1-11) x A91965 (1-7475)
HOFFMANN LA ROCHE (CH)
Location/Qualifiers
1. 7475
                                                                                                                                                                                                                                                                                                                                                                                                                                                  /organism-"unknown"
1916 c 2039 q
                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AR207294.1 GI:21506162
                                                                                                              3.53
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           3.53
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unclassified.
                                                                                                                                                                                                                                                                                                                                                                Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                 Best Local Similarity:
                                                                      ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1853 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Unknown.
                                                                                                                                                                                                                                                                                                                                         Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unknown
                                                                    1853
                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                     Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                      Alignment Scores:
                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                                                                                                                                                            KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                            source
                                                                                                                                                                                                                                                                                                                                                                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DEFINITION
                                                                  BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                            AR031992
LOCUS
                                                                                                                                                                                                                                                                                             DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM
                                                                                                                 Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                          TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                 PEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AR207294
            FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ò
```

```
HCJ238800 9033 bp RNA linear VRL 18-AUG-1999
Hepatitis C virus type 1b complete genome, isolate NC1.
AJ28800
AJ238800. GI:5748510
Complete genome: core protein; glycoprotein E1; glycoprotein E2;
NS2 proteinase; NS3 proteinase/Helicase; NS3/4A proteinase
cofactor; NS4b protein; NS5A phosphoprotein; NS5B RNA dependant RNA
Hepatitis C virus type 1b
Hepatitis C virus type 1b
Viruses; SSRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /codon_start=1
/product="polyprotein"
/product="polyprotein"
/product="polyprotein"
/protein_id="CAB53095.1"
/db_xref="G16"CAB53095.1"
/db_xref="G16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AKNTIGITSLFSPGSSQKIQLVNTNGSWHINRTALNCNDSLNTGFLAALFYVHKFNSS
GCPERMASCSPIDAFAQGWGPITYNESHSSDQRPYCWHYAPRPCGIVPAAQVCGPVYC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FTPSPVVVGTTDRFGVPTYSMCENETDVLLLNNTRPPQGNWFGCTWMNSTGFTKTCGG
PPCNIGGIGSKTLTCPTDCFRKHPEATYTKCGSGPWLLPRCLVHYPYRCTVNF
TIFKVRMYVGGVEHRLEAACNWTRGERCNLEDRDRSELSPLLLSTTEMOVLPCSFTTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (10-MAY-1999) Bartenschlager R., Institute for Virology, Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67, 55131 Mainz, GERMANY Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ሯ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Koch,J.O. and Bartenschlager,R.
and Abalation of hepatitis C virus NS5A hyperphosphorylation
nonstructural proteins NS3, NS4A, and NS4B
J. virol. 73 (9), 7138-7146 (1999)
                                                                                                                                                                                                                                                    7475
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /organism-"Hepatitis C virus type 1b"
                                                                                                                                                                                                                                                                                                         Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels:
                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /virion
/mol_type="genomic RNA"
/strain="RB"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /isolate="NC1"
/db_xref="taxon:31647"
/country="Germany"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-26 (1-11) x AR207294 (1-7475)
/organism="unknown"
1916 c 2039 g
                                                          б
```

```
RIGITREYTPGERPSGHPDSSVLCECTDGCANTELTTAATSYRLRAYLINTPGELPVCO
LILEPRESPYTGCLTHIDAHFLGGTVGAGDNPPTLAYOATVCARAQAPPSHDGWC
LILERAPTLHGPTPLLYRIGAYONEVTTHPITKYIMACHSADLEVVYSTWVVGGVLA
ALAAYCLTTGSVVIVGRIVLSGKPAIIPDREVLYREPDEMEECASHLPYIEQGMGLAS
PKOKAAGLOTATKQARAAPVDESKRTLLENGWAKHWHTSGIOYLAGLSTLPGN
PAIASLAWYTASITSPLTYOHTLLFNILGGWVAAQLAPPRAASAFVGAGIAGAAVGSI
GLGKVLVDILAGYGAGVAGALVAFKVNSGEMPSTEDLVNLLPAILSPGALVVCVVCAA
RICHWGRGGGAVOWMNRLIAFASKROMISPHIHVPESDAARVTQLISSSLTTYOLLK
RICHWINDECSTPCSGSWLRDWWNICTVIDFKTWLOSKLLPRLPGVPFFSGORGYY
GVWRGDGIMQTTCPCGAQIIGHVKNGSMRIFGPRTCSNTWHGTFPINAYTTGPCTPSP
                                                                                                                                       DSKGSLLSPRPVSYLKGSSGGPLLCPSGHAVG IFRANCTRGVAKAVDFVPVESMETT
MRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAGGYKVI.VLNPSVAATL
SFGAYMSKAHGIDPNIRTGVRTTTTGAPITYSTYGKFLADGGCSGGAYDIIICDECHS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APNYSRALMRVAAEEYVEVTRYGDFHYVTGMTTDNVKCPCQYPALEFFTEVDGYRLHR
YAPACKPLLREEVTFLVGLNQYLVGSQLPCEPEPDVAVLTSMLTDPSHITAETAKRLL
ARGSPPSLASSSASQLSAPSLKATCTTRHDSPDADLIEANLLMRQEMGGNITRVESEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KVV ILDSFEPLRAEEDEREVSVPAETLRRSRKFPRAMPTWARPDYNPPLLESWKDPDY
VPPVVHGCPLPPAKAPPTPPPRKRTVVLSESTVSSALAELATKTFGSSESSAADSGT
ATAPPDQFSDDGDAGSDVGSYSSMPPLEGEPGDPDLSDGSWSTVSEEAGEDVVCCSMS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVFPDIGYRVCEKWALYD
VSTLLQDAWGSSYGFQYSPOGRVEFLVNTWKSKRCPMGFAYDTRCPDSTYTENDIRV
EESIYQCCDLAPEARQAIRSLTELY IGGPLTNSKGQNCGYRRCRASGVLTTSCGNTL
TCYLKRAAACRARALOOCTMIVGGDDLVVICESAGTQEDEASLRAFTEAMTRYSAPPG
                                                                                                                                                                                                                            TDSTTILGIGTVLDQAETAGARLVVLATATPPGSVTVPHPNIEEVALSSTGEVPFYGK
AIPIETIKGGRHLIFCHSKKKCDELAAKLSGLGLNAVAYYRGLDVSVIPTSGDVIVVA
TDALATGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            YTWTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSASLROKKVTFDELQVLD
DHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKDVRNLSSKAVSHI
LOVWIPPLNVRGGRDAVILLTCAIHPELIFTITKILLAILGPLMVLQAGITKVPYFVR
                           AHGLIRACMIVRKVAGGHYVQMALMKLAALTGTYVYDHI,TPLRDWAHAGLRDLAVAVE
                                                                                    AYSQOTRGLLGCIITSLTGRDRNQVEGEVQMVSTATQSFLATCVNGVCWTVYHGAGSK
TLAGPKGPITOMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DPPRPEYDLELITSCSSNVSVAHDASGKRVYYLLTRDPTTPLARAAWETARHTPVNSWL
GNIIMYAPTLWARMILMTHFPSILLAQEQLEKALDCQIYGACYSIEPLDLPQIIQRLH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSQGGRAATCĞKYL
FNWAVRTKLKLTPIPAASQLDLSSWFVAGYSGGDYYRSLSRARPRWFMWCLLLLSVGV
                                                         PVVFSDMETKVITWGPDTAACGDIILGLPVSARRGREIHLGPADSLEGQGWRLLAHII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   'product="NS3 proteinase/helicase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /product="glycoprotein El"
1150, .2238
/gene="E2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /product-"glycoprotein E2"
2239. .2427
/gene-"p7"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /product="NS2 proteinase"
3079. .4971
/gene="NS3"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /product-"core protein"
574. :1149
/gene="El"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /product-"NS4B protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /product="p7 peptide"
2428. 3078
/gene="NS2"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /gene-"E2"
2239. "...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'gene="core"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5134. .5916
/gene="NS4B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   l. .573
/gene="core"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /gene="NS2"
3079. .4971
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /gene="p7"
2428. .3078
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GIYLLPNR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1. .573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            gene
```

```
Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients with hepatocellular carcinoma: the 'progression score' revisited Hepatol. Res. 20 (2), 161-171 (2001)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VRL 22-AUG-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis C virus
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Takahashi,K., Iwata,K., Matsumoto,M., Matsumoto,H., Nakao,K.,
Hatahara,T., Ohta,Y., Kanai,K., Maruo,H., Baba,K., Hijikata,M. and
Mishiro,S.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Direct Submission
Submitted (19-SEP-2000) Shunji Mishiro, Toshiba General Hospital,
Submitted of Medical Sciences, 6-3-22 Higashi Oh-1, Shinagawa ku,
Department of Medical Sciences, 6-3-22 Higashi Oh-1, Shinagawa ku,
Tokyo 140-8522, Japan (E-mail:shunji.mishiro@po.toshiba.co.jp,
Tel:81-3-3764-8981, Fax:81-3-3764-8992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nbv4yV92 9344 bp RNA linear V) Hepatitis C virus gene for polyprotein, complete cds, isolate:HCVT145.
                                                                                                                                                                                                                                                                      /product="NS5B RNA dependant RNA polymerase"
2731 c 2574 g 1903 t
4972. .5134
/gene="NS4A"
/product="NS3/4A proteinase cofactor"
5134. .5916
/gene="NS4B"
5917. .7256
                                                                                                                                                                                                                                                                                                                                                    9033
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                       5917. .7256
/gene-"NS5A"
5917. .7256
/gene-"NS5A
/product-"NS5A phosphoprotein"
7257. .9030
/gene-"NS5B"
/gene-"NS5B"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GlySerValValileValGlyArgileValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Hepatitis C virus"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /specific_nost=human"
/db_xref="taxon:11103"
/tissue_type='serum'
/note='HCV-genotype: 1b'
<1. .10
                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-26 (1-11) x HCJ238800 (1-9033)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /mol_type="genomic RNA"
/isolate="HCVT145"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AB049092
AB049092.1 GI:11559450
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (bases 1 to 9344)
                                                                                                                                                                                                                                                                                                                                                    4.33
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                      Best Local Similarity:
                                                                                                                                                                                                                                                                                       1825 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mishiro, S
                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                       mat_peptide
                                          mat_peptide
                                                                                                                                                          mat_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DEFINITION
                                                                                                                                                                                                                                                                                       BASE COUNT
                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5'UTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
             gene
                                                                                                                           gene
                                                                                                                                                                                                          gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 7
AB049092
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE
                                                                                                                                                                                                                                                                                                      ORIGIN
```

/note="partial"

VRL 02-AUG-1993

CDS

```
HEPCRNA 9456 bp RNA linear VRL 03-FEB-1999 Hepatitis C virus RNA, complete genome sequence. D10934.1 G1:471116
C; E; NS1/E2; NS2; NS3; NS4; NS5; polyprotein precursor. Hepatitis C virus Hepatitis C virus Viruses; serNa positive-strand viruses, no DNA stage; Flaviviridae; Hepativis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Wang, Y., Okamoto, H., Tsuda, F., Nagayama, R., Tao, Q.M. and Mishiro, S. Prevalence, genotypes, and an isolate (HC-C2) of hepatitis C virus in Chinese patients with liver disease J. Med. Virol. 40 (3), 254-260 (1993)
                                                                                                                                                                                                                                                                     Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
                                                                                                                                                                                                                                                                                                                                                                                                                                     Original source text: Hepatitis C virus (individual_isolate China, strain Hebel) (library: NC1-18) CDNA to genomic RNA.

Location/Qualifiers

1. 9400
//organism="Hepatitis C virus"
//mol_type="genomic RNA"
/strain="Hebel"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      These data kindly submitted in computer readable form by: Hiroaki
                                                                                                                                                                                                                                                                                                                                                                     from mainland of
                                                                                                    lincar
                                                                                                                                                                                                                                                                                                                 1 (bases 1 to 9400)
Bi.S.-1., Bai.X.-H., Margolis,H.S. and Liu,C.-B.
Complete hepatitis C virus genome cDNA isolated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9400
111
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative:
                                                                                            HPCCGENOM 9400 bp ss-RNA LO2836. China virus complete genome. L02836.1 G1:329739 complete genome: viral genome. Hepatitis C virus Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5342 GGCAGCGTGGTCATTGTGGGCCAGGATCGTCCTG 5374
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2650 g 2034 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            US-09-965-594-26 (1-11) x HPCCGENOM (1-9400)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /db_xref="taxon:11103"
/tissue_lib="NC1-18"
355. .9384
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /isolate="China"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (bases 1 to 9456)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Jichi Medical School
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.52
49.00
100.00$
100.00$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Immunology Division
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2796 c
                                                                                                                                                                                                                                                                                                                                                                                                                   Unpublished (1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Kawachi-gun
Tochigi 329-04
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ۵
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           J. Med. V
93359897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1920
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            8394876
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Okamoto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    source
                                                                                            LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                             ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
MEDLINE
PUBMED
                                              RESULT 8
HPCCGENOM
                                                                                                                                                                                                                                                                                                                                           AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                     COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HPCRNA
                                                                                                                                                                                                                     SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                        CLEARS A LATER TO SHARE THE TOTAL TH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSVWKDLLEDTETPIOTTIMANNEVECVOPEKGGRKPALIYPPDLGYRVCEKHALIY
VVSTLLEQAWGSSYGFQYSFGQRVEFLYNAMKSRKRPMGFAYDTRCFESTYTENDIRY
BESIYQCCDLAPEARLAIRSLTFELYVGGPHTNSKGONGGYRRCRASGVLTTSGGNLT
TCYLKASAACRAAKLRCTMLVGGDULVVICESAGTQEDBAASIRVFFBAMTRYSAPG
DPPQPEYDLELITGCSSNVSYAHDANNKRVYYLTRDPTIPLSRAAWETARHTPVNSWL
GLIMYPALHARMILMTHFSILLAGOQLKRALOOLOGARISISLEDLDQVIQRLH
GLSAFSLHTYSPEGIRVASGLKKGVPPLRWWHRARSVRAKLLGGGGRALIGKYL
FNWAVRTKLKLTPIPAASRLDLSSWFVAGYSGGDIYHSVSRARPFRWFLLLLSVGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATAPPDQPLDDGGTGSDVESYSSMPPLEGEPGDPDLSDGSWSTVSEEGGEDVVCCSMS
YTWTGALITPCAAEESKLPINALSNSLLRHHNLYYATTSRSASQRQKKVTFDRLQVLD
DHYRDVLKEWKAKASTVKARLLSVEEACKLTPPHSARSKFGYGAKDVRSLSSKAVNHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9344
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2655 g 1958
                                              /product-"polyprotein"
/protein_id-"BAB18805.1"
/db_xref-"GI:11559451"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.49
49.00
100.008
100.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2842 C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GIYLLPNR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
Best Local Similarity:
Query Match:
```

0285-44-2111 x3334 0285-44-1557.

Phone:

1 GlySerValValIleValGlyArgIleValLeu 11

US-09-965-594-26 (1-11) x AB049092 (1-9344)

1889

BASE COUNT

NIBIRC

Alignment Scores: Pred. No.: Japan Fax: ø

source

FEATURES

5'UTR CDS

```
Conservative:
Mismatches:
Indels:
        Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /db_xref="taxon:11103"
/country="Russia"
                                                                                                                                                                                                                                                        US-09-965-594-26 (1-11) x HPCRNA (1-9456)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AF176573.1 GI:5738246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /isolate-"lb"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepacivirus.
1 (bases 1 to 9600)
Mokhonov, V.V., Samok)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    with acute hepatitis
49.00
100.008
100.008
100.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Prilipov, A.G.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Unpublished
                                           Percent Similarity:
Best Local Similarity:
                                                                                                                              Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5'UTR
CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
AF176573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                I MOREAALENLYVILNDEYLINGARIYOTENENLINGARIYOTEN KONTONINGARIYOTEN K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DHLEFWESVFTGLTHIDAHFLSOTKOAGDNFPYLVAYOATVCARAAPPPSWDOMWKC
LIRLKPTLHGPTPLLYRLGAVONEVILTHPITKYIMACWSADLEVYTSTWVLVGGVLA
LAAAYCLITGSVYUVGRILSGRPAVIPDREVULYOEFDEMEEGCSHLFPTEOGMQLAE
OFKOKLGLLOIATROABAAPVESKWRALETFWARHWMYTSGVOYLAGISTLPGN
PAIASLAAFTASVTSPLTTQSTLLFNILGGWVAAQLAPPSAASAFVGACIAGSAVGSI
GLGKVLVDILAGYGAGVAGALVAFKINGSGFPSAEDIVNLLPAILSPGALVGVCVAA
RLHOWINEDCSTPGSGSWLEDIWMWICSVLDFKTWLGSKLLPRLPGVPFFSCORGY
RLHOWINEDCSTPGSGSWLEDIWMWICSVLDFKTWLGSKLLPRLPGVPFFSCORGY
GVWRGDGIMQTTCPCGAQITGHVKNGSMRIVGPKTCSNTWHGTFPINAYTTGPCTPSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SRPSWGPNDPRRESRNLGKVIDTLTCGFADLWGYIPLVGAPLGGAARALAHGVRALED
GVWYATGNLPGCSFSIFLLALLSCLITPASAFEYRNAGGYYTVTNDCSNSSIVYEAAD
HMHNPGCVPCVRENNSSRCWYALTPLARNASVPTTIRRHVDLLVGAAALCSAMY
VGDLCGSVELVSQLFFSPRRHETVQPCCNCSIYPGHTGHRAMADWANNSPTTALVY
SQLLRIPQAVMDMVAGAHWGVLAGLAYYSMVGNWAKVLIVLLLFAGVDGNTYVTGGAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          YAPACKPVLREEVDEGVGLNOYPVGSOLPCEPBPDVAVLTSMITDPSHITAEAAKRRI.
ARGSPPSLASSSASOLSAPSLKATCTTHHDSPDADLIEANLLWRGEMGGNITRVESEN
KVVILDSFEPLRAEEDEREVSVAAEILRRTRRFPPAMPVWARPDYNPPLLESWKDPDY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VPPVVHGCPLPPTKAPPIPPPRKRTVVLTESTVSSALAELATKTFGSSGSSAVDSGT
ATAPPDQTSNDGDRESDAESYSSMPPLEGEPGDPDLSDGSWSTVSEEASGDVVCCSMS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TCYLKASAACRAAKLQDCTMLVNGDDLVVICESAGTQEDAASLRVFTEAMTRYSAPPG
DPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARAAWETARHTPVNSWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GNIIMYAPTLMARMILMTHFFSILLAQEQLEKALECQIYCACYSIEPLDLPQIIERLH
GLSAFSLHSYSPGEINRVASCLRKLGVPPLRVWRHRARSVRAKLLSQGGRAATCGKYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ARGASGITSLFSRGPSQKIQLVNTNGSWHINRTALNCNDSFNTGFLAALFYAHRFNSS
GCPERMASCRSIDKFDQGWGPITYYQGDSPDQRPYCWHYPPRRCGIVPASEVCGPVYC
FTPSPVVVGTTDRLGVPTYNWGENETDVLLLNNTRPPQGNWFGCTWMNTTGFTKTCGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PPCNIGGAGNNTLICPTDCFRKHPEATYTKCGSGPWLIPRCLVDYPYRLMHYPCTVNF
TIFKVRMYVGGVEHRLDAACNWTRGERCALEDRDRSELSPLLLSTTEWQILPCSFTTL
PALSTGLIHLHRNIVDVQYLYGIGSAVVSFAIKWEYVLLLFLLLADARVCACLMMALL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           YTWTGALITPCAAEESKLPINPLSNSLLRHHNMVYATTSRSASLRQKKVTFDRLQVLD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  rsvwkdlledtetpidttimaknevfcvqpekggrkparlivfpdlgvrvcekmalyd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /translation-"mstnpkpqrktkrntnrrpqdvkfpgggjvggvyllpprrgpr
gvratrktsersqprgrrqpipkarrpegramaqpgypwplygnegmgwagwllsprg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       apny sralwrvaaeeyvevtrvgdfhy itgmttonik cpcqvpapefftevdgvrlhr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EESIYOCCDLAPEAKOAIKSLIERLYIGGPLINSKGONCGYRRCRASGVLTTSCGNTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FNWAVRTKLKLTPIPAASRLDLSGWFVAGYGGGDIYHSLSRARPRWFMLCLLLLSVGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DHYRDVLKEMKAKASTVKAKLLSIEEACKLTPPHSAKSKFGYGAKDVRNLSSKAVNHI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VVSTLPQAVMGSSYGFQYSPGQRVEFLVNTWKSKKCPMGFSYDTRCFDSTVTENDIR
                                                                                                                                                                                                                                                                             342...9374
/note-"C, E, NS1/E2, NS3, NS4 and NS5"
/codon_start-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a T-stretch of 45 bp."
                                                                                                                                                                                                                                                                                                                                                                                                                    /product-"polyprotein precursor"
/protein_id="BAA01728.1"
/db_xref="GI:471117"
                                                                                  'organism="Hepatitis C virus"
                                                                                                                         /mol_type="genomic RNA"
/isolate="HC-C2"
/db_xref-"taxon:11103"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2668 g
    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9375. .9456 /note="with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2811 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GIYLLPNR 9375. .94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1909 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
```

Pred. No.:

BASE COUNT

ORIGIN

```
Hepatitis C virus polyprotein precursor, gene, complete cds. AF176573
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /ttaislation-"mstnpkportikrntnrrpoddyrpgggivggyildprrpprl
gvratryssersoprgrropipkarhpbgrtwaopgypwplygnbglgwagwllsprg
srpswgptdprrrsnigkyidtltgfadlmgyiplugaplggvaralahgyrvled
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GVNYATGNLPGCSFSIFLLALLSCLTIPASAYEVRNVSGVYHVTNDCSNSSIVYBAAD
VIMHTPGCVPCVQDGNTSRCWVALTFTLAARNASVPVTAIRRHVDLLVGTAAFCSAMY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SQLLRIPQAIVDMVAGAHWGVLAGLAYSMVGNHAKVMIVLLLFAGVDGTTHTTGGAA
ARATQGFTSFFSLGPSQKIQLINTNGSWHINRTALNCNDSLQTGFLAALFYTYRFNAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GCPERMASCRPIDKFDÖGMĞPITYAEPDSSDQRPYCHHYAPRPCGIVPASQVCGPVYC
FYSSPVVQTTDRFGVPTYTYGENETIVLLLURIYRPPLGMRYGGTWMRYGTGTKTCGG
PPCNIGGAGMTTLITOPTDCFRKHPENTYTKGGSGPWITPRCIVDYPXEMHYPANVI
TIFKVRMYVGGVEHRLNAACNWTRGERCDLEDRDRSELSPLLLSTTEMQVLPCSFTTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AYSQOTRGLLGCIITSLTGRDKMQVEGEVQVVSTATQSFLATCVNGVC#TVYHGAGTK
TLAGPRGPITQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRRG
DSRGSLLSPRPISYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVESNETT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VGDLCGSVFPVSQLFTFSPRRHQTVQDCNCSIYPGHISGHRMAWDMMMWSPTAALVV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PALSTGLIHLHQNIVDVQYLYGIGSAVIPFAIKWEYVLLLFLLLADARVCACLWMMLL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ia<u>oaeaa</u>lenivvlnaasvagahgiisfivffcaawyikgrivpgaayalygvwplli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LLLALPPRAYAMDREMAASCGGAVFVGLVLLTLSPHYKVFLARLIWWSQYFTTRAEAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            <u>Lovwypplnirggrdaiillmcavhselifditkfllailgplmmfqaritrvpyfvr</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AQGLI RACMLVRKAAGGHY IQMALMKLAALSGTYVYDHLTPLQDWAHAGLRDLAAAVE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PVVFSDMETKVITWGADTAACGDIISGLPVSARRGREILLGPADGLEEQGWRLLAPIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Molecular cloning HCV Russian isolate 1b from the serum of patient
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            342. 9374
//note="cleaved into C, E1, E2, p7, NS2, NS3, NS4a, NS4b,
NS5a, and NS5b proteins"
/codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 (bases 1 to 9600)
Mokhonov, V.V., Samokhvalov, E.I., Novikov, D.V., Shatalov, A.G. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ., Samokhvalov, E.I., Novikov, D.V., Shatalov, A.G. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Direct Submission
Submission
Submitted (09-AUG-1999) Molecular Genetics, Ivanovsky Virology
Institute, Gamaleya Str., 16, Moscow 123098, Russia
Location/Qualifiers
1. 9600
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
/strain="274933RU"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /note="isolated from acute hepatitis patient serum"
40000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /product-"polyprotein precursor"
/protein_id-"AAD50312.1"
/db_xref-"G1:5738247"
```

```
Percent Similarity:
Best Local Similarity:
                                                            Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5'UTR
CDS
                                                                                                                                                                                                                                                                                                                                                                                                              DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE
PUBMED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                                                                           RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                    AF268278
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SOURCE
                                                                                                                                                                                                                                                               g
                                                                                                                                                                 THE PURE THE POTE TO STATE TO 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HSVÄKDLLEDTETPIDTTIMAKNEVFCVQPEKGGRKPARLIVPPDLGVRVCEKMALYD
VSTLEQAWGSATGROYSPGGRVDFFLYAMKRKCPMGFSPTDTRCPBTTEBIRV
VSTLEQAWGSATGROYSPGGRVDFFLYAMKRKCPMGFSPTDTRCPBTTTEBIRV
VESITOCCDLAPEAGRAIRSLTERLYVGGPLTNSKGONGCYRRCRASGVLTTSGCWIT
TCYLKATAACRAAKLRDCTLLVNGDDLVVICESAGTGEDAASLRVFTEAMTRYSAPPG
TCYLKATAACRAAKLRDCTLLVNGDDLVVICESAGTGEDAASLRVFTEAMTRYSAPPG
GNIIMYAPTUMARWILLATHFFSILLFQCGLEKALDCOITGAYYSIEPLDLPOITGRLH
GLSAFSLHSYSPGEINRVASCLRKLGVPPLRAMMRRARNVRAKLLSGGGRAATGGKYI.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PAT 20-APR-2002
                                                         TDSTTILGIGTVLDOAETAGARLVVLATATPPGSVTVPHPNIEEIALSNTGEIPFYGK
TPIEIIKGGRHLIFCHSKKCDELARAKLSGGLANAYXTRAGLDVSVTPPSGNVVVVA
TDALATGETGDFDSVLOKTFYTQTVDFSLDPFFIIETTYPQDAVSRSQRRGRERR
RRGIYRFVTPGERPSGAFDSSVLCECYDAGCAMYELTPAETSVRLRAYLNTPGLPVCQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FNWAVRTKLKLTPIPAASOLDLSGWFVAGYGGGDIYHSLSRARPRWPMLCLLLLFVGV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 (bases 1 to 12734)
Hong, Z., Lai, W.C.H. and Lau, J.Y.N.
Hong, Z., Lai, W.C.H. and Lau, J.Y.N.
Hepatitis C virus proteased-dependent chimeric pestivirus
Patent: US 6326137-A 1 04-DEC-2001;
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12734
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               9600
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GlySerValValileValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5. .9600
2887 c 2696 g 2100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 1 from patent US 6326137.
ARI79057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-26 (1-11) x AF176573 (1-9600)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /organism="unknown"
2604 c 3295 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR179057.1 GI:20220612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               4.62
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GIYLLPNR"
9375. .96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6.27
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unclassified
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4032 a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 11
AR179057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DRIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
```

```
LOUGHEWING HOW THE LEGGEN TO A THE MAINTEN THE MAD A THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pestivirus type 1
Pestivirus type 1
Viruses: ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Viruses: Lo 12734)
Lai, V.C., Ednor, W., Skelton, A., Ingravallo, P., Vassilev, V.,
Donis, R.O., Hong, Z. and Lau, J.Y.
Generation and characterization of a hepatitis C virus NS3
protease-dependent bovine viral diarrhea virus
J. Virol. 74 (14), 6339-6347 (2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2 (bases 1 to 12734)
Lai, V.C.H. and Hong, Z.
Direct Submission
Submitted (16-MAY-2000) Antiviral Therapy, Schering-Plough Research Institute, 2015 Galloping Hill Road, Kenilworth, NJ 07033-0539, USA Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SYSDTKEEGATKKKTOKPDKLERGKMKIVPKESEKDSKTKPPDATIVVEGVKYOVRKK
GKTKSKNTQDGLYHNKNKPQESRKKLEKALLAMAIIAIVLFQYTMGENITQMNLQDNG
TEGIQRAMFQRGVNRSLHGIMPEKICTGVPSHLATDIELKTIHGMADASEKTNYTCCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TVVLIVIGLIIARRDPTIVPLVTIMAALRVTELTHQPGVDIAVAVMTITLLMVSVYTD
YFRYRKWLQCILSLVSGVPLIRSLIYLGRIEMPEVTIPNWRPLTLILLYLISTTIVTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /translation="MELNTNEGSGSVVIVGRIVLSGSGSITACAOOTRGLLGCKITSL
TARRNAW DEGEVOIVSTATOPTELATOL NGVCWTYVTBAGFTIASPRGPVIONTNUD
QDIVGWPAPQGSRSITPCTGSSDLYLVTRHANVIPVRRRGDSRGSLLSPRFISTISTE
SGGGPLLCPAGHAVGLFRAAVCTRGVAKAVDFIPVENLETTTRSGSGADTEDVVCCSM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LTLDFMYYMHRKVIEEISGGTNIISRLVÄALIELMWSMEEEESKGLKRFYLLSGRLRN
LIIKHKVRNETVASWYGEEEVYGMPKIMTIIKASTLSKSRHCIICTVCEGREWKGGTC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VRL 12-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IDYTRVDSIYDVDESGEGVYLFPSRQKAQGNFSILLPLIKATLISCVSSKWQLIYMSY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
     0000
          Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Pestivirus type 1"
/mol_type="genomic RNA"
/db_xref="taxon:11099"
1 .385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RNA
                                                         Mismatches:
                                                                                                                                                                                                                                                                                                                                                          1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pestivirus type 1, complete genome. AF268278
                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /product-"polyprotein"
/protein_id-"AAF82566.1"
/db_xref-"G1:9049957"
                                                                                                                                                                                                                                                           JS-09-965-594-26 (1-11) x ARI79057 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          12734 bp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /codon_start=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AF268278.1 GI:9049956
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10864644
```

```
Single-chain recombinant complexes of hepatitis C virus NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AR145221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown.
                                                                                                                   16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                          Alignment Scores:
                                                                                                                                                                                                                                                                    Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BASE COUNT
ORIGIN
                                                                                                               BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 15
AR145197
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 14
AR14521
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AUTHORS
                                     JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURES
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOURCE
                                                                                                                                                                                                                Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
                                                                                                                                                                                                                                                            HYXKCYTAK IDY SKRALLATDKURWEVEHGVITRIJAKRYTGVGFNGATIGDEPHRALV
ERDCATITKNTVQFLKMKKGCAFTYDLISNLTRIJELVHRNNLEEKEI PTATVTWL
ATTEVNEDDSOVOTKPVLGETVEDPVVDINLQPEVQVDTSEVGITIGRETLMTTGVTP
VLEKVEPDBSDNONSVKIGLDEGNYPGPGIQTHITEEIHNDARPTIMILGSTNSIS
NRAKTARNINLYTGNDEEIRDLMAAGRALVVALRDVDFELGEMVDFKGTFLDREALE
ALSLGQPKPKQVTKEAVRNIEDGKKDVEIPNWFASDDPVFLEVALKNDKYYLVGDVGE
                                                                                           GARCYVLNPEAVNISGSKGAVVHLQKTGGEFTCVTASGTPAFFDLKNLKGWSGLPIFE
ASSGRVVGRVKVGKNEESKPTKIMSG1QTVSKNTADLTEMVKKITSMNRGDFKQ1TLA
TGAGKTTELPRAVIEEIGRHKRVLVLIPLRAAAESVYQYMRLKHPSISFNLRIGDMKE
                                                                                                                                                                                                        LPDLDTVIDTGLKCEKRVRVSSKIPFIVTGLKRMAVTVGEQÄQRRGRVGRVGRYKPGRYYR
SQETATGSKDYHYDLLQAQRYGIEDGINVTKSFREMNYDWSLYEEDSLLITQLEILINN
LLISEDLPAAVKNIMARTDHPEPIQLAYNSYEVQVPVLFPKIRNGEVTDTYENYSFLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VKDTVIREHNKWILKKIRFOGNLNTKKMLNPGKLSEOLDREGRKRNIYNHQIGTIMSS
AGIRLEKLPIVRAQTDIKTFHEAIRDKIDKSENRONPELHNKLLEIFHTIAQPTLKHT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              YGEVTWEOLEAGINRGAAGFLEKKNIGEVLDSEKHLVEOLVRDLKAGRKIKYYETAI
SKURKRUSVEDDOMGADLVEKRRPKIOYDEAKTRAITKVWYMWYKOPPVIOOTEGK
TPLENIFDKVRKEWBENEPVAKYEDTKAMOYDOVISKDLOLICIGEIOKYYYKKEHHKEI
DTITDHMTEVPVITADGEVYIRNGQRGSGOPDISAGNSMLNVLTMMYARCESIGVPYK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IEFCSHTPVPVRWSDNTSSHMAGRDTAVILSKWATRLDSSGERGTTAYEKAVAFSFLL.
MYSWNPLVRRICLLVLSQQPETDPSKHATYYYKGDPIGAYKDVIGRNLSELKRTGFEK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PAT 08-AUG-2001
                                                                                                                                                   GDMATGITYASYGYFCQMPQPKLRAAWVEYSYIFLDEYHCATPEQLAIIGKIHRFSES
IRVVAMTATPAGSVTTTGQKHPIEEFIAPEVMKGEDLGSQFLDIAGLKIPVDEMKGNM
LVFVPTRNMAVEVAKKLKAKGYNSGYYSGEDPANLRVVTSQSPYVIVATNAIESGVT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VKDOAKALGATDOTRIIKEVGSRTYAMKLSSWFLOASNKOMSLTPLFEELLLRCPPAT
KSNKGHMASAYQLAQGNWEPLGCGVHLGTIPARRVKIHPYEAYLKLKDFIEEEEKKPR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SFNRVARIHVCGDDGFLITEKGLGLKFANKGMQILHEAGKPQKITEGEKMKVAYRFED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LANLNLSLSTLGIWTKHTSKRIIQDCVAIGKEEGNWLVNADRLISSKTGHLYIPDKGF
                                                                         rgwaythoggissydhytagkdlivcdsmgrtryvcqsnnrltdeteygyktdsgcpd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TLOGKHYEQLQLRTETNPVMGVGTERYKLGPIVNLLLRRLKILLMTAVGVSS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Unknown.
Unclassified.
1 (bases 1 to 75)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 others
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12734
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       413 GGTAGTGTTGTTATTGTTGGTAGAATTGTTTTA 445
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 48 from patent US 6211338.
AR145217
AR145217.1 GI:15107084
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x AF268278 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3293 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6.27
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12509. .127
1 2608 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AR145217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3'UTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
AR145217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ACCESSION
VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
```

```
PAT 08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        08-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 (bases 1 to 75)
Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Single-chaln recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 55 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Universitie.

I (bases 1 to 78)

Malcolm, B.A., Taremi, S.Shane., Weber, P.C. and Yao, N.
Sightle-chaln recombinant complexes of hepatitis C virus NS3
protease and NS4A cofactor peptide
Patent: US 6211338-A 26 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PAT
                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        linear
                                                                                                                                 75
10
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0001000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA
                                                                                                                                                                                                                                                                          DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ب
                                                                         ų
protease and NS4A cofactor peptide
Patent: US 6211338-A 48 03-APR-2001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 26 from patent US 6211338.
AR145197
                                                                                                                                                                                                                                                                                                                                                           Sequence 55 from patent DS 621[338.
AR145221
AR145221.1 GI:15107088
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                         30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-26 (1-11) x AR145221 (1-75)
                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x AR145217 (1-75)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="unknown"
12 c 16 a
                              Location/Qualifiers
                                                         /organism="unknown"
14 c 15 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualiflers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AR145197.1 GI:15107064
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    48.00
100.00%
90.91%
97.96%
                                                                                                                              0.0408
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Unclassified.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
```

	78	10	7	0	0	0
31 t	ngth:	tches:	nservative:	smatches:	dels:	Gaps:
18 g	re Le	æ	ပ္ပ	Mi	ä	Ga
11 c	0.0426	48.00	100.00%	90.916	896.76	9
183	cores:		ilarity:	Similarity:	•	
BASE COUNT	Alignment So	Score:	Percent Sim	Best Local	Query Match	DB:
	18 3 11 c 18 g	Scores: 0.0426 Length:	Scores: 0.0426 Length: 48.00 Matches:	11 c 18 g 31 t 0.0426 Length: 48.00 Matches: 100.00% Conservative:	Scores: 0.0426 Length: 48.00 Matches: 100.004 Conservative: 51milarity: 90.91% Mismatches:	0.0426 Length: 48.00 Matches: 100.00% Conservative: 90.91% Mismatches: 97.96% Indels:

Search completed: August 31, 2003, 00:46:50 Job time : 151.976 secs

S G

US-09-965-594-26 (1-11) x AR145197 (1-78)

Run on:

```
Hepatitis C virus
HCV NS4A DNA fragm
HCV NS4A DNA fragm
HCV Polynucleotide
HCV Polynucleotide
HCV Peptide C14-1.
HCV I chimeric epi
HCV I chimeric epi
DNA encoding non-A
NON-A, NON-B Hepat
HCV antigen. Synt
DNA encoding Hepat
HCV antigen. Synt
DNA associated wit
ANTI-VIRAL Synt
HCV NS4A-NS3 compl
                                                                                                                                                                  Chimeric BVDV/HCV
HCV NS4A-NS3 compl
                                                                                                                           Hepatitis C virus
Hepatitis C virus
HCV NS2-NS4 clone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protease; viral replication; chronic liver disease; liver cancer; primer; ss.
                                                                       Description
and is derived by analysis of the total score distribution
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        HCV NS4A-NS3 fusion protease oligonucleotide #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ALIGNMENTS
                            SUMMARIES
                                                                                                                                                                     ABA95615
                                                                                                                                                                                                                                                                                                                                                                                                                                AAQ25753
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       BP
                                                                                                AAA73336 standard; DNA; 68
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                    Length
                                        Query
Match I
                                                                                                                                       Hepatitis; NS3
liver failure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Synthetic
                                                                                                                                                                                                                                                                              Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA73336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73336
                                                                                                                                                                                                                                                                              Result
                                                                    Š
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (without alignments)
2906.924 Million cell updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N. Geneseq_19Jun03:*

| SIDSI/gcgdata/geneseqn-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1981.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1986.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1980.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA1990.DAT:*
| SIDSI/gcgdata/geneseqq-embl/NA2001.DAT:*
                                                                                          August 30, 2003, 19:13:57; Search time 10.2149 Seconds
                                                                                                                                                                                                                                                                                                                   5105512
           GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.
                                                                  nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                                                                                 of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                        2552756 seqs, 1349719017 residues
                                                                                                                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                                                                                                      0.5
0.5
7.0
                                                                                                                                                                                                                                    Ygapext
Fgapext
                                                                                                                                                                                                                                                              , Delext
                                                                                                                                                                                                                                                                                                                                           length: 0
length: 2000000000
                                                                                                                                                 US-09-965-594-26
49
                                                                                                                                                                             1 GSVVIVGRIVL 11
                                                                                                                                                                                                      BLOSUM62
Xgapop 10.0 , X
Ygapop 10.0 , Y
Fgapop 6.0 , F
Delop 6.0 , D
                                                                                                                                                                                                                                                                                                                                             sed
                                                                                                                                                                Perfect score:
                                                                                                                                                                                                         Scoring table:
                                                                                                                                                                                                                                                                                                                   Total number
                                                                                                                                                                                                                                                                                                                                           Minimum DB s
Maximum DB s
                                                                  OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database :
                                                                                                                                                                                Sednence:
                                                                                                                                                                                                                                                                                         Searched:
```

~

```
Score:
                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                                                                              ದ್ದ
                                                                                                                                                                                      linker segments which were used in the creation of a fusion molecule of the coding sequences for Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                               present sequence is one of two oligonuclectides coding for NS4A and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protease; viral replication; chronic liver disease;
                                                                                                                                                                                                                                                                                                                                  0000
                                                                          Goldfarb V;
                                                                                                                                                                                                                                                                                                      other;
                                                                                                                                                                                                                                                                                                                                          Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HCV NS4A-NS3 fusion protease oligonucleotide #2.
                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 68 BP; 20 A; 14 C; 17 G; 17 T; 0
                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               liver failure; liver cancer; primer; ss
                                                                          Weinheimer S, Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x AAA73336 (1-68)
                                                                                                                                                           Example 2; Page 22; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              8
                                                        (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                    protease; inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (BRIM ) BRISTOL-MYERS SQUIBB
                    2000WO-US00345.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        06-JAN-2000; 2000WO-US00345
                                      99US-0115271
                                                                                                                                                                                                                                                                                                                               0.106
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73337 standard; DNA; 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (first entry)
                                                                                             WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200040707-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hepatitis; NS3
                                                                          Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           08-JAN-1999;
                    06-JAN-2000;
                                                                                                                                                                                                                                                                                                                        Alignment Scores:
                                      08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13-JUL-2000
13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA73337;
                                                                                                                                                                                                                                                                                                                                                                     Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                           Score:
g
                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
The present sequence is one of two oligonuclectides coding for NS4A and linker segments which were used in the creation of a fusion molecule of the coding sequences for Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is useful as HCV can lead to chronic liver disease such as cirrhosis, liver failure and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this
                                                                                                             Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                type, as well as enabling structural studies of the protease and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hepatitis C virus NS4A-NS3 fusion protease coding sequence #1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /*tag= a
/product= "NS3-NS4A fusion protein"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Goldfarb V;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 71 BP; 17 A; 18 C; 15 G; 21 T; 0 other;
     Zhang Y, Goldfarb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      60 GGATCCGTTGTTATCGTCGCCCGTATAGTACTG 28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2hang Y,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (1-71)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                 Example 2; Page 23; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (BRIM ) BRISTOL-MYERS SQUIBB CO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-26 (1-11) x AAA73337
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAA73328 standard; DNA; 588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  990S-0115271.
Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0.111
49.00
100.00%
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               19-DEC-2000 (first entry)
                                                      WPI; 2000-465976/40.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wittekind M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAA73328;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 3
```

```
8
                                                                                                                   The present sequence is the coding sequence for a fusion protein created using the Hepatitis C virus (HCV) MS3 and MS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV unfection. This is useful as HCV can lead the present invention concerns a number of NS3 mutants and NS3·NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and
                                  Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix 0 amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hepatitis; NS3 protease; viral replication; chronic liver disease;
liver failure; liver cancer; mutant; muteln; ds.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Hepatitis C virus NS4A-NS3 fusion protease coding sequence #2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           /*tag= a
/product= "NS4A-NS3 fusion protein #2"
                                                                                                                                                                                                                                                                                                        0
0
0
0
0
0
                                                                                                                                                                                                                                                                    Sequence 588 BP; 97 A; 183 C; 153 G; 155 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Goldfarb V;
                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                13 GGTTCCGTTGTTATCGTCGCCCGTATAGTACTG 45
                                                                                                                                                                                                                                                                                                                                                                                                                    1 GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                             Mismatches:
                                                                                                                                                                                                                                                                                                                                                        Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zhang Y,
                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-26 (1-11) x AAA73328 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                               Disclosure; Fig 10; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (BRIM ) BRISTOL-MYERS SQUIBB CO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAA73329 standard; DNA; 588 BP
                                                                                                                                                                                                                                              protease:inhibitor complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wittekind M, Weinheimer S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              06-JAN-2000; 2000WO-US00345
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99US-0115271
                                                                                                                                                                                                                                                                                                       1.21
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-DEC-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WPI; 2000-465976/40.
P-PSDB; AAB15220.
2000-465976/40
                                                                                                                                                                                                                                                                                                                   Score:
Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hepatitis C virus.
            P-PSDB; AAB15212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200040707-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       08-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       13-JUL-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAA73329;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAA7332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            a
                                                                                                                                                                                                                                                                                                                                                                                                                     8
```

```
fusion protein created using the Hepatitis C virus (HCV) NS3 and NS4A protease enzymes. These proteins are both essential for the replication of the virus, acting to cleave its replicative proteins from the replication polyprotein produced from the HCV genome. Inhibitors of the two proteins should be effective as antiviral treatments of HCV infection. This is should be effective as antiviral treatments of HCV infection. This is mutants and liver cancer. The present invention concerns a number of NS3 mutants and NS3-NS4A fusion proteins which can be used to identify inhibitors of this type, as well as enabling structural studies of the protease and protease:inhibitor complexes. The protein produced from this sequence contains the alpha-helix0-1 variant.
Modified hepatitis C virus (HCV) NS3 protease comprising at least 1 substitution of a hydrophobic alpha-helix \theta amino acid to a hydrophilic amino acid, useful for screening inhibitors that may treat hepatitis C
                                                                                                                                                                                                                                          present sequence is the coding sequence for a mutated version of a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Н
С
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Clone; polypeptide; NS2-NS4; Hepatitis C; Virus; HCV; serum;
transcriptase; cDNA; primer; allele; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Takahashi K;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 588 BP; 103 A; 180 C; 156 G; 149 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             588
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sek1 M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-965-594-26 (1-11) x AAA73329 (1-588)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Honda Y, Murakami T,
                                                                                                                                                                          Claim 26; Fig 12; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (MITU ) MITSUBISHI KASEI CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91JP-0172794.
91JP-0287008.
91JP-0332329.
92JP-0099957.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92EP-0109812.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          91JP-0139268
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1.21
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (updated)
(first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HCV NS2-NS4 clone N13-1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAQ32477 standard; DNA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9..608
/*tag=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       07-0CT-1991;
16-DEC-1991;
20-APR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-MAR-2003
26-APR-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11-JUN-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11-JUN-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12-301-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16-DEC-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EP518313-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hayashi N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAQ32477;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAQ3247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Y CONTRACTOR OF THE STATE OF TH
```

Teranishi Y;

```
The sequences given in AAQ32472-82 and AAQ32442 are various clones which encode the NS2-NS4 regions of the Hepatitis C Virus (HCV) gene of the invention. These sequences were isolated from the serum of a patient suffering from hepatitis C (HC). The isolated RNA sequences were converted into cDNA using transcriptase in the presence of one of the primer sequences given in AAQ3251964. The sequences were then amplified using primer pairs. The cDNA sequences isolated represent different alleles of the same region of the HCV gene. Sequence comparisons of these clones showed that it is possible for a AAQ32436.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nucleic acid construct encoding chimeric Hepatitis C Virus (HCV) pestivirus genome where the Npro protease gene is replaced with NS3 protease gene, useful for in vivo screening of compounds which inhibit
                                                   New hepatitis C virus gene and its encoded protein - used for diagnosing and vaccinating against hepatitis C virus infections {\sf det}
                                                                                                                                                                                                                                                                                                             Sequence 618 BP; 135 A; 184 C; 176 G; 123 T; 0 other
                                                                                                                                                                                                                                                                                                                                                                      0000
                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pestivirus; Npro; protease; NS3; screening; ds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                   (Updated on 25-MAR-2003 to correct PN field.)
                                                                                                                                                                                                                                                                                                                                                                                                               Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Columns 17-28; 20pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Chimeric - Bovine viral diarrhea virus.
Chimeric - Hepatitis C virus.
                                                                                                Disclosure; Page 125-26; 305pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-26 (1-11) x AAQ32477 (1-618)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Chimeric BVDV/HCV NS3-wt sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABA95615 standard; DNA; 12734 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0344456
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Lai VCH, Lau JYN;
                                                                                                                                                                                                                                                                                                                                                        1.28
49.00
100.008
100.008
100.008
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2002-121103/16.
         WPI; 1992-417213/51.
P-PSDB; AAR29846.
                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          infection
                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-JUN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US6326137-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     04-DEC-2001.
                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Example 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hong Z,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

```
chimeric Repatitis C virus (HCV)-pestivirus genome. The construct comprises a pestivirus genome where a Npro pestivirus protease gene is replaced with a gene encoding a functional HCV NS3 protease. Furthermore, each junction site recognised by the Npro protease is replaced with a junction site recognised by the HCV NS3 protease. The construct is useful for screening compounds that inhibit HCV in vivo by inhibiting HCV protease, where screening may be in cell culture or in an animal model. The present sequence is a chimmeric clone of BVDV (bovine viral diarrhea virus)/HCV NS3-wt, which was used to illustrate the present invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the ATPase activity of NS3. The covalent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HCV: hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
             present invention relates to a nucleic acid construct encoding a
                                                                                                                                                                                                       Sequence 12734 BP; 4032 A; 2604 C; 3295 G; 2803 T; 0 other;
                                                                                                                                                                                                                                                           12734
11
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complex construction primer SEQ ID NO:48.
                                                                                                                                                                                                                                                                            Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                          1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                  Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-26 (1-11) x ABA95615 (1-12734)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 1; Page 32; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  980S-0094331.
970S-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528
                                                                                                                                                                                                                                                           39.4
49.00
100.00%
100.00%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAX80310 standard; DNA; 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                   Best Local Similarity:
                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             W09928482-A2
                                                                                                                                                                                                                                              Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80310;
                                                                                                                                                                                                                                                                                                                                    Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAX80310
                                                                                                                                                                                                                                                                                 Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ò
```

S

Alignment Scores:

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS3 serine protease domain. Where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                          HCV NS4A-NS3 complex construction primer SEQ ID NO:55.
                                                                                                                               27
0
0
0
0
0
0
                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
Indels:
Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 75 BP; 16 A; 12 C; 16 G; 31 T; 0 other;
                                                                             Sequence 75 BP; 16 A; 14 C; 15 G; 30 T; 0 other;
                                                                                                                                                                                                                                                                                     GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                       US-09-965-594-26 (1-11) x AAX80310 (1-75)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Example 1; Page 33; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  98WO-US24528.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98US-0094331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97US-0067315
                                                                                                                           0.186
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                               AAX80314 standard; DNA; 75
                                                                                                                                                                                                                                                                                                                                                                                                                            07-SEP-1999 (first entry)
                                  present sequence represency NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Synthetic.
Hepatitis C virus.
                                                                                                                                          Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-NOV-1997;
                                                                                                           Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                               AAX80314;
                                                                                                                                                                                                                                                                                                  13
                                                                                                                                                                                       Query Match:
DB:
                                                                                                                                                                                                                                                                                                                              ឧ
                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                  a
```

```
NS4A-NS3 complex contraining a central hydrophobic domain of native HCV NS4A-NS3 complex a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A serine protease domain, where the hydrophobic domain of native HCV NS4A serine protease domain, where the hydrophobic domains of the HCV NS3 protease domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by MNR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                                                                                                                                                                                                                                                                     HCV; hepatitis C virus; single chain recombinant complex; linker; NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease; hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      present invention describes a covalent hepatitis C virus (HCV)
                                                                                                                                                                                                                                                                                            HCV NS4A-NS3 complex construction primer SEQ ID NO:42.
 75
10
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              78
10
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 78 BP; 16 A; 14 C; 17 G; 31 T; 0 other;
                           Conservative:
Mismatches:
                                                                                                                        Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Yao
                                                                                                US-09-965-594-26 (1-11) x AAX80314 (1-75)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weber PC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Example 1; Page 31; 211pp; English.
                                                                                                                                                                                                           AAX80306 standard; DNA; 78 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0094331.
0.186
48.00
100.00%
90.91%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             0.195
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                  07-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                              Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                           Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                      WO9928482-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                10-JUN-1999
                                                                                                                                                                                                                                       AAX80306;
                                                       Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ..
0
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      The
                                                                                                                                                                               RESULT 9
                                                                                                                                                                                                AAX80306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred.
                 Score:
                                                                                                                                                                                                                                       d
```

9

RESULT 10

```
US-09-965-594-26 (1-11) x AAX80338 (1-96)
                                                                                                                                                                                                                                                                                                                                                                                                           S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 2; Page 43; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                            Weber
                           AAX80338 standard; DNA; 96 BP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80340 standard; DNA; 96 BP
                                                                                                                                                                                                                                                                                                        98WO-US24528.
                                                                                                                                                                                                                                                                                                                                  98US-0094331.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0.246
48.00
100.00%
90.91%
97.96%
                                                                                     (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                           Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                            (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 1999-385385/32
                                                                                                                                                                                                      Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                               W09928482-A2
                                                                                                                                                                                                                                                                                                      24-NOV-1998;
                                                                                                                                                                                                                                                                                                                                  28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                                                     07-SEP-1999
                                                                                                                                                                                                                                                                           10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                         Malcolm BA,
                                                        AAX80338;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match:
RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAX80340
               aax80338
                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         The present invention describes a covalent hepatitis C virus (HCV) NSAA-NS3 complex comparison a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 perclease domain. The covalent NS4A NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the Arpase activity of NS3 complexes are more soluble, stable and active than the
                                                                                                                                                                                                                                           HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                  HCV NS4A-NS3 complex construction primer SEQ ID NO:26.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             78
10
10
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 78 BP; 18 A; 11 C; 18 G; 31 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-965-594-26 (1-11) x AAX80290 (1-78)
          JS-09-965-594-26 (1-11) x AAX80306 (1-78)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Example 1; Page 26; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Weber PC,
                                                                                                                                ВР
                                                                                                                                                                                                                                                                                                                                                                                                                                   98US-0094331,
97US-0067315.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0.195
48.00
100.00%
90.91%
97.96%
                                                                                                                              AAX80290 standard: DNA; 78
                                                                                                                                                                                       (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (SCHE ) SCHERING CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 1999-385385/32
                                                                                                                                                                                                                                                                                                                 Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                               W09928482-A2
                                                                                                                                                                                     07-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                        24 - NOV - 1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                         0-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Malcolm BA,
                                                                                                                                                                                                                                                                                                     Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13
                                                                                                                                                          AAX80290;
```

Query Match: Best Local

Score:

```
The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A peptide, a linker, and an HCV NS3 estine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3 protease domain. The covalent MS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity, the helicase activity and the APPase activity of NS3. The covalent non-covalent protease peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                    HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
HCV NS4A-NS3 complex construction primer SEQ ID NO:87.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Seguence 96 BP; 21 A; 17 C; 21 G; 37 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Yao
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-SEP-1999 (first entry)
```

```
Hepatitis C Virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      21-NOV-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ABX1006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The present invention describes a covalent hepatitis C virus (HCV) NS4A-NS3 complex comprising a central hydrophobic domain of native HCV NS4A-NS4 peptide, a linker, and an HCV NS3 serine protease domain, where the hydrophobic domain of native HCV NS4A peptide is tethered by the linker to the amino terminus of the HCV NS3A peptide is tethered by the linker to the amino terminus of the HCV NS3A protease domain. The covalent NS4A-NS3 complexes are useful for structural determination and determination of mode of binding of HCV inhibitors by NMR spectroscopy. They can also be used for detecting inhibitors of the protease activity and the ATPase activity of NS3. The covalent NS4A-NS3 complexes are more soluble, stable and active than the non-covalent protease-peptide complexes previously available. The present sequence represents a primer used in the construction of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Double stranded RNA; dsRNA; RNAI; RNA inhibition; cytostatic; virucide; proteozoadide; gene expression; antisense; tumour; infection; Plasmodium; virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus; Hepatitis C virus; human papalloma virus; gene; ds.
                                                                   HCV; hepatitis C virus; single chain recombinant complex; linker;
NS3 protease; NS4A cofactor; NS4A-NS3 complex; serine protease;
hydrophobic domain; covalent complex; detection; inhibitor; primer; ss.
                       HCV NS4A-NS3 complex construction primer SEQ ID NO:89.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hepatitis C virus NS4A polynucleotide SEQ ID NO 129.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     96
10
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 96 BP; 21 A; 17 C; 21 G; 37 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Weber PC, Yao N;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New hepatitis C virus covalent complexes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-26 (1-11) x AAX80340 (1-96)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Example 2; Page 44; 211pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABZ35821 standard; DNA; 161 BP
                                                                                                                                                                                                                                                                                                                                                                980S-0094331.
                                                                                                                                                                                                                                                                                                                    98WO-US24528
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0.246
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07-FEB-2003 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Malcolm BA, Taremi SS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         HCV NS4A-NS3 complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                     (SCHE ) SCHERING CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 1999-385385/32.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                 Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                             28-JUL-1998;
28-NOV-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                    24-NOV-1998;
                                                                                                                                                                                                                              W09928482-A2
                                                                                                                                                                                                                                                                         10-JUN-1999
                                                                                                                                                           Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AB235821;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AB235821
LID AB2
LID AB2
AC AB2
XX
XX
XX
DDT 07-
DX
XX
DE Hep
XX
CM DOU
XX
C
셤
```

```
The invention relates to inhibiting expression of a target gene in a cell by introducing at least two oligoribonuclectides (dsrNaI and II), both with a double-stranded (ds) structure of at most 49 sequential nucleotide pairs. At least part of one strand (Si, S2) of the ds structures in each of dsrNaI and II are complementary to regions in the target gene. The method uses antisense inhibition of gene expression using double stranded RNA inhibition (RNAI). The method is particularly used to treat tumours or infections, especially by plasmodium or viruses/viroids (pathogenic on humans, animals or plants). The method provides more effective inhibition of expression than known methods using a single dsRNA, even at very low concentrations. When dsRNA has at least one unpaired nucleotide at the end, stability (and thus effective concentration in the cell) is improved and efficiency can be increased further by pretreating the cells with interferon. The present sequence is that of a target DNA of the
                                                                                                                                                                                                                                                                                                                                                                                                 Inhibiting expression of target genes, useful e.g. for treating tumors, by introducing into cells two double-stranded RNAs that are complementary to the target
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        interferon; oncogene; cytokine; Id; developmental;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161
0 0 0
0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  55 G; 34 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                     Hadwiger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-26 (1-11) x ABZ35821 (1-161)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HCV NS4A DNA fragment SEQ ID 129.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Claim 13; Page 87; 100pp; German.
                                                                                                                                                                                                                                                                                     Rost S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABX10064 standard; DNA; 161 BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence 161 BP; 32 A; 40 C;
                                                                                                            09-JAN-2001; 2001DE-1000588.
                                                                                                                                                                     09-JAN-2001; 2001DE-1000588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                0.442
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23-JAN-2003 (first entry)
                                                                                                                                                                                                                                                                                     Kreutzer R, Limmer S,
                                                                                                                                                                                                                              (RIBO-) RIBOPHARMA AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Oligoribonucleotide;
prion; inhibition; ds
                                                                                                                                                                                                                                                                                                                                           WPI; 2002-683450/74.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hepatitis c virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DE10100587-C1.
DE10100588-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Alignment Scores:
                                                         18-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABX10064;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               invention
```

œ

07-DEC-2001; 2001DE-1060151.

```
This invention describes a novel method for inhibiting expression of a target gene by introducing into the cell that contains the target gene at least one oligoribonuclectide (GRNNI) that has a double-stranded (ds) structure of not more than 49 consecutive nucleotides (nt), where at least a segment of one strand of the ds structure is complementary with the target gene and the cells are treated with interferon before introduction of dsRNAI. The method is used to inhibit expression of target genes, particularly oncogenes, cytckine genes, Id (not defined) protein genes; developmental or prior genes corpuses or virois pathogenic organisms (particularly plasmodia) or in viruses or virois (pathogenic in humans, animals or plants). Treating the cells with interferon greatly increases the extent to which dsRNA can inhibit expression of the target genes, and the effect is even greater when dsRNA are modified to increase their stability, ABX09936-ABX10075 represent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RNA inhibition; dsRNA1; gene expression inhibitor; oncogene; cytostatic; virucide; protozoacide; gene; ds.
                                                                                                                                                                           Inhibiting expression of target genes, e.g. oncogencs, in cells, by introduction of complementary double-stranded oligoribonucleotide, after treating the cell with interferon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    gene fragments used to illustrate the method of the invention
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   161
10
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 161 BP; 32 A; 40 C; 55 G; 34 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Hadwiger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus NS4A DNA SEQ ID NO 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x ABX10064 (1-161)
                                                                                                                                                                                                                                         Disclosure; Page 92; 98pp; German.
                                                                                                             Rost S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         09-JAN-2001; 2001DE-1000586.
26-OCT-2001; 2001DE-1055280.
29-NOV-2001; 2001DE-1058411.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          09-JAN-2002; 2002WO-EP00152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0.442
48.00
100.00%
90.91%
97.96%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABV78245 standard; DNA; 161
                 09-JAN-2001; 2001DE-1000587
                                                09-JAN-2001; 2001DE-1000587
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-NOV-2002 (first entry)
                                                                                                             Kreutzer R, Limmer S,
                                                                                (RIBO-) RIBOPHARMA AG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
                                                                                                                                              WPI; 2002-742209/81
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hepatitis C virus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO200255693-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-JUL-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABV78245;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ABV78245
a
```

```
The invention relates to inhibiting expression of a target gene (I) in a cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded structure of at most 49 consecutive bases. At least part of one strand (asi) of dsRNA1 is complementary to (I) and at least part of one strand has an overhang of 1-4 nucleotides. The method is used to inhibit the expression of a wide range of genes, e.g. oncognese, cytokine genes etc. In humans, also genes in plasmodium or in viruses or viroids that are pathogenic for humans, animals or plants. Introducing an overhang into dsRNA greatly increases effectiveness for inhibiting que expression, both in vivo and in vitro and also increases stability and thus the effective concentration inside the cell. The present sequence is that of
                                                                                                                                                          the
                                                                                                                                   Inhibiting expression of target gene, useful e.g. for inhibiting oncogenes, by administering double-stranded RNA complementary to target and having an overhang -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     161
10
1
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 161 BP; 32 A; 40 C; 55 G; 34 T; 0 other;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches:
                                                                    Hadwiger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Search completed: August 30, 2003, 19:48:23 Job time : 12.2149 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-965-594-26 (1-11) x ABV78245 (1-161)
                                                                                                                                                                                                            Claim 10; Page 190; 203pp; German.
                                                                                                                                                                                                                                                                                                                                                                                                                                               gene related to the invention.
                                                                  Rost S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     0.442
48.00
100.00%
90.91%
97.96%
                                                                  ŝ
                                (RIBO-) RIBOPHARMA AG
                                                                                                     WPI; 2002-590671/63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Alignment Scores:
                                                                    Kreutzer R,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score:
õ
```

GenCore version 5.1.6 Copyright (c) 1993 - 2003 Compugen Ltd.

```
August 30, 2003, 19:20:43; Search time 106.667 Seconds (without alignments) 2506.388 Million cell updates/sec
OM protein - nucleic search, using frame_plus_p2n model
                                                                                                                                                                                                                                                              22781392 seqs, 12152238056 residues
                                                                                                                                                                               0.5
7.0
7.0
                                                                                                                                                            BLOSUM62
Xgapop 10.0 , Xgapext
Ygapop 10.0 , Ygapext
Fgapop 6.0 , Fgapext
Delop 6.0 , Delext
                                                                                                US-09-965-594-26
49
1 GSVVIVGRIVL 11
                                                                                                Title:
Perfect score:
Sequence:
                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                                                                Searched:
                                Run on:
```

45562784 hits satisfying chosen parameters: Total number of

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

EST:* Database

gb_est3:*
gb_est4:*
gb_est5:*
em_estfun:*
em_estom:*
em_gss_hum:* em_gss_mus: *
em_gss_pro: *
em_gss_rod: *
em_gss_rod: * em_gss_inv:* em_gss_pln:* em_gss_fun:* em_gss_mam:* em_gss_vrt:* em_estpl:* em_estro:* em_htc:* gb_est2:* gb_htc:* em_esthum: em_estba:* em_estin: * em_estmu:* em_estov:* gb_est1:*

em_gss_vrl:* gb_gssl:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

u	"	
Ġ	٥	
۲	4	
٥	Ľ	
4	ζ	
è	Ę	
2	Ę	
	2	

		. مع				
Result No. So	core	Query	Length	80	ID	cription
	45	1.8	23,	6	W580	W580055 RC1-HT03
0	43	87.8	618	14	582	28563 wdilc.p
m	43	Ļ.	664	m	BW300987	987 BW30098
	.	٠,	709	6 0 (BW306038	038 BW30603
n w	4 4	٠ ٧	(7/	7	BW1/6544	5544 BW176544
۰,	4 4	85.7	646	2 -	BG59081	0 0
ထ	42		658	(1)	BO121075	075 ESTADES
σ	42	Š	777	0	BG600305	305 EST50520
10	42	S	1035	m	BU718639	SJM2CLF1
11	41	83.7	222		AU183349	49 AU183349
c 12	41	۳.	487	\sim	B1882963	BI882963
13	41	3.	502	0	BE556909	909 fk92d02
14	41	ω.	564	4	CB417286	7286
~	41	e,	584	N	BI981053	1053 fu38h04
_	41	۳,	597	~	BI840872	0872 fq39q12
7	41	ω.	603	\sim	BM185756	3756 fv71h12
_	41	щ.	607	Ē	BQ093551	3551 fz48e06
~	41	ω,	620	7	BI982002	2002 fu58b11
~	41	ä	629	0	BE016243	5243 fk68h06
c 21	41	e,	641	0	BE016238	238 fk68g10
~	41	Э.	641	0	BE201405	405 fk92b09
23	41		660	4	CD392599	599 Gm_ck11
c 24	41	ë.	676		AW019436	36 felle05.
01 (41	83.7	796		AF122168	AF122168
~ (4	'n.	882	m 1	BQ734390	390 AGENCOU
c 27	41	m i	988	(1)	BI759235	235
~ (41	m,	1569	σ	CC250893	1893 СН261-1
7 1	.	i,	315		AV211292	292 AV211292
30	9	∹.	398	σ,	CNS03X0L	414
. س	9	ή.	540	0	BC403954	3954 6024198
32	40	ή,	667	4	CB030392	392 TGESTZY
T	0	ή,	693	28	115	1158 oeg89
m	0.4	ä	771	13	206	5063 Ae4-382A
32	40	≓,	791	29	206	5061 Pan t
Μ,	40	ä	1280	29	250	2509 Pan trog
37	40	.	1657	. 10	376	3760 60267303
38	40	.i.	1976	11	595	5952 Mus musc
36	30	ģ	116	10	926	E069569 RC2-BT03
0.7	36	6	135	10	E06956	9567 RC2-BT
4.1	39	ó	170	10	7113	1133 RC3-BT05
42	39	œ.	178	10	E06973	3730 RC2-BT03
43	39	e.	195	10	E07110	71108 RC3-BT05
O 44	33	Ġ	234		254	12254 AV012254
4	39	φ.	251	10	B44	B441428 BB
					ALIGNMENTS	
RESULT 1						
	AWS	80055				MRNA 1 TOPAT TECH 16-MAR-2000
DEFINITION		-HT0375	-130100	-011	HT0375	abiens CDNA mRNA s
ACCESSION		AW580055				

EST.

Homo sapiens (human)

Homo sapiens

Homo sapiens

Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 234)

HCGP http://www.ludwig.org.br/ORESTES. AW580055.1 GI:7255104 VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS

~

```
235
43.00
90.918
72.738
87.768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ciona intestinalis
Ciona intestinalis
                Seq primer: M13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Aliqnment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 3
BW300987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             | Control of the property of the profile of the pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             wdilc.pk004.f22 wdilc Triticum aestivum cDNA clone wdilc.pk004.f22 5' end, mRNA sequence. CA728563 GA728563 GI:25450552 EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticane: Tritica
                                                                                                                                                                                                                                                                  Fax: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl-RC1&t2-RC1-HT0375-
130100-011-905&t3-2000-01-13&t4-1)
Seq primer: puc lB forward
High quality sequence start: 28
High quality sequence store: 216.
                                         Unpublished
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 E. I. DuPont de Nemours and Company
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104, USA
Fax: 302-631-2602
Email: Scott.V.Tingey@USA.dupont.com
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        234
0000
        FAPESP/LICR Human Cancer Genome Project
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EST.
Triticum aestivum (bread wheat)
Triticum aestivum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-09-965-594-26 (1-11) x AW580055 (1-234)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Contact: Scott V. Tingey
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26.4
45.00
100.00%
72.73%
91.84%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Crop Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
Best Local Similarity:
Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 2
CA728563/c
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
JOURNAL
COMMENT
TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   à
```

```
BW300987 Nori Satch unpublished cDNA library, neural complex Ciona intestinalis cDNA clone cinc020906 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.

1 (bases 1 to 664)

Satou, Y., Shin-i, T., Kohara, Y. and Satoh, N.

Expressed genes in Ciona intestinalis (2002c)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /tissue_type="neural complex"
/clone_lib="Nori Satoh unpublished cDNA library, neural
complex"
                                                                                                                                                                                                 618
2
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6664
0 0 0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Contact: Nor1 Satch
Department of Zoology
Kyoto University
Sayo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satch@sscidian.zool.kyoto-u.ac.jp.
1. .664
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism-"Clona intestinalis"
                                                                                                                                                                                                                                                                                                            1, .618 // /organism-"Triticum aestivum"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      183 t
                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cinc020g06"
                                                                                                                                                                                                                                                                                       US-09-965-594-26 (1-11) x CA728563 (1-618)
cocation/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    135 g
                                                                                                                                                                                                                                                                                                                                                                                                                           BW300987
BW300987.1 GI:24881598
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      259
43.00
100.00%
90.00%
87.76%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      145 c
```

```
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
Location/Qualifiers
                                                                                                                                                                                                                                                                                   Conservative:
Mismatches:
       Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
                                                                                                                                                                                                                                                     Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                       indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /organism="Homo saplens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        High quality sequence start: 3
High quality sequence stop: 150.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-26 (1-11) x BW176544 (1-725)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BE069681
BE069681.1 GI:8414331
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens (human)
Homo sapiens
                                                                                                                                                                                                                                                                  43.00
100.00%
90.00%
87.76%
                                                                                                                                                                                                                                                                           Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10737800
                                                                                                                                                                                           195
                                                                                                                                                                                                                                     Alignment Scores:
Pred. No.:
                                                                                    source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         source
                                                                                                                                                                                       BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         JOURNAL
MEDLINE
PUBMED
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 6
BE069681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AUTHORS
                                                                  FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VERSION
                                                                                                                                                                                                                                                                     Score:
                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                        BW306038 Nori Satch unpublished cDNA library, heart Clona intestinalis cDNA clone clht012f15 5', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BW176544 Tinear EST 04-NOV-2002 BW176544 Nori Satoh unpublished cDNA library, heart Ciona intestinalis cDNA clone rciht012f15 3', mRNA sequence.
                                                                                                                                                                                                                                          Eukaryotan Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona. (1 (bases 1 to 709) Satou, Y., Shin i, T., Kohara, Y. and Satoh, N. Expressed genes in Ciona intestinalis (2002c)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota: Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Cionidae; Ciona.

1 (bases 1 to 725)
Satou, Y., Siin-i, T., Kohara, Y. and Satoh, N.
Expressed genes in Ciona intestinalis (2002c)
                                                                                                                                                                                                                                                                                                                                                                                                                                                           1...709
./organism="Ciona intestinalis"
./organism="Ciona intestinalis"
./mol_type="mRNA"
./db_xref="contolis"
./clone="ciht012f15"
./clone="type="heart"
./clone=lib="hori Satch unpublished CDNA library, heart"
a 156 c 155 g 199 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                709
1
0
0
                                                                                                                                                                                                                                                                                                                                     Contact: Nori Satch
Department of Zoology
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satch@ascidian.zool.kyoto-u.ac.jp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
Matches:
Conservative:
Mismatches:
Indels:
                                               307 GGAAGTGTGGTGGTTGGAAGGATTGTT 336
                                    10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 GlySerValValIleValGlyArgIleVal 10
                               1 GlySerValValIleValGlyArgIleVal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-26 (1-11) x BW306038 (1-709)
JS-09-965-594-26 (1-11) x BW300987 (1-664)
                                                                                                                                                                      BW306038.1 GI:24886649
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BW176544
BW176544.1 GI:24566468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Contact: Nori Satoh
Department of Zoology
Kyoto University
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               283
43.00
100.00%
90.00%
87.76%
                                                                                                                                                                                                                  Ciona intestinalis
Ciona intestinalis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ciona intestinalis
Ciona intestinalis
                                                                                                                                                                                                                                                                                                                           Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match:
DB:
                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 5
BW176544/c
                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pred. No.:
                                                                                                                                                                                                                                                                                                        TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              JOURNAL
                                                                                                                                                                                                                                                                             REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   REFERENCE
                                                                                          RESULT 4
BW306038
                                                                                                                                                                                                                                                                                                                                                                                                                                                FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score:
                                                                                                                           Locus
```

```
BE069681 150400-014-c02 BT0389 Homo sapiens cDNA, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutherla; Primates; Catarrhini; Hominidae; Homo.
1 (bases I to 150)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Franstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Tel: +55-11-2704922
Frax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/qethtml2.pl?tl=&t2=RC2-BT0389-120Seq primer: puc 18 forward
1. 725
/organism="Ciona intestinalis"
/mol_type="mmNA"
/db_xref="taxon:7719"
/clone="rciht012115"
/tissue_type="heart"
/clone=lib=Nori Satch unpublished cDNA library, heart"
a 168 c 155 g 207 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Shotgun sequencing of the human transcriptome with ORF expressed sequence tags Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000) 20202663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
```

```
/clone="STMCY29"
/tissue_type="mixed tissues"
/lab_host="sol"mixed potato tissues"
/clone_lib="mixed potato tissues"
/clone=lib="mixed potato tissues"
/clone=lib="mixed potato tissues"
/clone=lib="mixed potato tissues"
/note="Vector: pBluesoript SK(-); Site_l: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora infestens-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BQ121075 658 bp mRNA linear EST 07-MAR-2003 EST606651 mixed potato tissues Solanum tuberosum cDNA clone STMEV29 5° end, mRNA sequence. BQ121075 G1:21920306
      this sequence is actually derived from Phytophthora rather than potato. \mbox{\tt l}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  on Apr. 17, 2002 this sequence version replaced gi:20173037. Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Solanum tuberosum (potato)
Solanum tuberosum
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Bukaryota; Viridiplantae; Streptophyta; Eudicotyledons; core eudicots;
asterids; lamiids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 658)
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
Karamycheva, S.A.
Generation of a set of potato cDNA clones for microarray analyses
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    658
11
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Location/Qualifiers
1. 658
Aolatsm-Solanum tuberosum"
Aolatype-"mRNA"
Cultivar-"Rennebec or Binjte"
Ab_xref-"taxon:4113"
                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative: Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                220 t
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels:
                                                                                                                                                                                                                                                                                                                       Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tubers, or roots."
97 c 182 q
                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-965-594-26 (1-11) x BG590884 (1-646)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    399
42.00
90.91%
81.82%
85.71%
                                                                                                                                                                                                                       42.00
90.91%
81.82%
85.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Seq primer: T3.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                Best Local Similarity:
                                                                        159 a
                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alignment Scores:
                                                                                                                                                                  Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                           Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred. No.:
                                                                     BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEFINITION
                                                                                                                                                                                                       . . ON .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 8
BQ121075
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   l. 646
/organism="Solanum tuberosum"
/wol_type="mRNA"
/cultivar="Expe="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="Bulfinfs"
/tissue_type="leaf"
/dev_stage="6 week old"
/dab_bost="SolR"
/clone_lib="P: infestans-challenged leaf"
/doce="vector: paluescript SK(-): Site_l: EcoRI; Site_2:
/note="vector: paluescript SK(-): Site_l: Docore Sporangia/ml): Leaf tissue was collected at 1, 2, 2, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR. Katahdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that
/clone_lib-"BT0389"
/note-"Organ: breast; Vector: pucl8; Site_l: SmaI; Site_2: SmaI; Amin:-library was made by cloning products derived from ORESTES FOR (U.S. Letters Patent application No. 196 ,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mrNA and cDNA amplification were performed under low stringency conditions."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BG590884 646 bp mRNA linear EST 07-MAR-2003 EST498726 P. Infestans-challenged leaf Solanum tuberosum cDNA clone BPLI6H5 5' sequence, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1 (bases 1 to 646)
2 handage 1, conditions of the following states of the foll
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotylodons; core eudicots;
asterids; lamiids; Solanales; Solanaceae; Solanum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
18mail: Potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
                                                                                                                                                                                                                                                                                                                                                                                   Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              http://genome.arizona.edu/orders/
Seg primer: M13F-R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-965-594-26 (1-11) x BE069681 (1-150)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BG590884.1 GI:13609024
                                                                                                                                                                                                                                                                                                                                                                                      55.3
42.00
90.91%
72.73%
85.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          source
                                                                                                                                                                                                                                                            BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ACCESSION
VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REFERENCE
AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT 7
BG590884
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
```

q ò

,

S

```
Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
Strigeidida: Schistosomatoidea; Schistosomatidae; Schistosoma.
1 (Bases I to 103)
Hu.W. Yan.O.; Shen.D. Liu.F.; Xu.X.; Zhu,Z.; Zhang,X.; Wang,J.;
Sun,J.; Xu,X.; Wang,Z.; Zeng,L.; Rong,Y.; Wu,X.; Ou,J.; Xu,Z.;
Huang,J.; Ma,Y.; Wang,S.; Mang,Z.; Xue,C.; Feng,Z.; Chen,Z. and Han
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidea; Cyp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AU183349 222 bp mRNA linear EST 30-JUL-200 AU183349 Cyprinus carpio head kidney stimulated by lipo-polysaccharide and concanavalin-A Cyprinus carpio cDNA clone H7, mRNA sequence.
                                                                                                                                                                                                                                                               Expressed sequence tags from male adults of Schistosoma japonicum
                                                                                                                                                                                                                                                                                          Unpublished
Contact: Zequang Han
Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Miyazaki University
1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan
Emali: m.sakalecc.miyazaki-u.ac.jp.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /organism-"Schistosoma japonicum"
/mol_type="mRNA"
/db_xref="taxon:6182"
/sex="male"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /tissue_type="Whole body"
/dev_stage="adult"
/lab_host="rabbits"
/clone_lib="sym"
13 c 311 g 310
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-965-594-26 (1-11) x BU718639 (1-1035)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cyprinus carpio (common carp)
Cyprinus carpio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
1. .1035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 311 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AU183349.1 GI:13919090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Contact: Masahiro Sakai
Faculty of Agriculture
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Email: hanzg@chgc.sh.cn
              Schistosoma japonicum
                                         Schistosoma japonicum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       42.00
90.91%
72.73%
85.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      11742760
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21607774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DEFINITION
                                   ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BASE COUNT
                                                                                                                                                                                                                                                               TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE
AUTHORS
TITLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE
PUBMED
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 11
AU183349
                                                                                                                         REFERENCE
                                                                                                                                                    AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
                                                                                                                                                                                                                                                                                                                           COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORIGIN
           SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /clone_lib="cSIS"
//note="Vector: pBluescript SK(-); Sitc_l: EcoR; Site_2:
Xho; various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26c in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        BU718639 10-FEB-2003 SJM Schistosoma japonicum cDNA, mRNA sequence. BU718639 GI:28325986 EST.
                                                                                                                                                                                                                               BG600305 777 bp mRNA linear EST 07-MAR-2003 EST505200 cSTS Solanum tuberosum cDNA clone cSTS28L13 5' sequence,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unpublished
Contact: Nobin Buell
The Institute for Genomic Research
9712 Medical Center Dr. Rockville, MD 20850, USA
Email: potato-array@tigr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 77)
van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B. Generations of ESTs from sprouting potato eyes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /tissue_type="Sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/iab_host="SOLR"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism-"Solanum tuberosum"
/mol_type-"mRNA"
/cultivar-"Kennebec"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-965-594-26 (1-11) x BG600305 (1-777)
US-09-965-594-26 (1-11) x RQ121075 (1-658)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /db_xref="taxon:4113"
/clone="cSTS28L13"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                        Solanum tuberosum (potato)
                                                                                                                                                                                                                                                                                                                                 BG600305.1 GI:13617441
EST.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       499
42.00
90.91%
81.82%
85.71%
                                                                                                                                                                                                                                                                                                                                                                                                                                   Solanum tuberosum
                                                                                                                                                                                                                                                                                       mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                             VERSION
KEYWORDS
SOURCE
ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DEFINITION
ACCESSION
VERSION
                                                                                                                                                                                                                                                         DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                  ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BU718639
LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AUTHORS
                                                                                                                                                                         RESULT 9
BG600305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score:
```

EST 30-JUL-2002

셤 ŝ

9

```
/organism-"Danio rerio"
/mol_type-"mRNA"
/db_xref-"taxon:7955"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                US-09-965-594-26 (1-11) x BI882963 (1-487)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /tissue_type="Fin"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Danio rerio (zebrafish)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BE556909.1 GI:9821399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           41.00
100.00%
63.64%
83.67%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                          144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                     BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ..
Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TITLE
JOURNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT 13
BE556909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LOCUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Contact: Stephen L. Johnson
Washington University School of Medicine.
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
7Tel: 314 286 1810
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Preparation: Ning Wu. CDNA Library Arrayed by: Steve
CDNA Library Arrayed by: Washington University Genome Sequencing
CONTACT CON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BI882963 487 bp mRNA linear EST 16-SEP-2002 fn01e01.x1 Zebrafish Research Genetics C32 fin Danio rerio cDNA clone IMAGE:4469065 3', mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    i (bases 1 to 487).
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy, S.L., Hiller,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood, K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,I., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                /clone="IMAGE:4469065"
/tissue_type="Fin"
/lab_host="GeneHogs (HS996, a phage-resistant isolate of
                                                                                                                                                                                                                      /clone="H7"
/clone_lib="Cyprinus carpio head kidney stimulated by lipo-polysaccharide and concanavalin-A"
/note="common name:common carp; stimulated by lipo-polysaccharide and concanavalin-A"
/note="common name:common carp; stimulated by lipo-polysaccharide and concanavalin-A"
/ 45 9 70 the state of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /clone_lib="Zebrafish Research Genetics C32 fin"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       222
7
7
0
0
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
Matches:
Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            42 GGGGCAGTGGTTGTCATGGGAAGAGTTGTACTG 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                        /organism="Cyprinus carpio"
/mol_type="mRNA"
/db_xref="taxon:7962"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WashU Zebrafish EST Project 1998
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           High quality sequence stop: 444.
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-965-594-26 (1-11) x AUI83349 (1-222)
    Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   information.
Seq primer: T7 from Gibco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BI882963.
BI882963.1 GI:16090234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Danio rerio (zebrafish)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           146
41.00
100.00%
63.64%
83.67%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1. .487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    and Wilson, R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Danio rerio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                   source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BI882963/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TITLE
JOURNAL
COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURES
FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               VERSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NIBING
```

```
Danio rerio
Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Butaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Actinopterygil: Neopterygil: Teleostei: Ostariophysi: Cypriniformes

1 (bases 1 to 502)
Clark,M. Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
'S., Hillier,L., Mucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
'K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BE556909 502 bp mRNA linear EST 30-AUG-2000 (K) 482402.yl Zebrafish Research Genetics C32 fin Danio rerio CDNA 5', mRNA sequence.
BE556909
/note-"Vector: pT7T3D-Pac with a modified polylinker;
Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was prepared
from zebrafish(C32) fin, and was then primed with a Not I
oligo(dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library is non-normalized. Library was
constructed by Ning Wu. NOTE: This clone is available
royalty-free through LLNL; contact the IMAGE Consortium
(info.inl.gov)for further information.

a 115 c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /close_llb="Zebrafish Research Genetics C32 fin"
/close="Vector: pT/T3D-Pac with a modified polylinker;
Site_l: EcoRI; Site_2: NotI; 1st strand cDNA was prepared
from zebrafish(C32) fin, and was then primed with a Not I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /lab host-"GeneHogs (HS996, a phage-resistant isolate of DH10B)"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              and Wilson, R.
Washu Zebrafish EST Project 1998
L Unpublished
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1810
Fax: 314 286 1810
Email: Zbrafish@watson.wustl.edu
cDNA Library Preparation: Ning Wu. cDNA Library Arrayed by:
Research Genetics. DNA Sequencing by: Washington University Geno
Sequencing Center Clone distribution: Research Genetics web
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     4 B 7 C 0 0 0 0 0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches:
Conservative:
Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         address: http://www.researchgenetics.com/
Seq primer: T3 ET from Amersham
High quality sequence stop: 466.
Location/qualiflers
```

```
and Wilson, R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    www.rzpd.de)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unpublished
                                                                                             Percent Similarity:
Best Local Similarity:
Query Match:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            159 a
                               Alignment Scores:
Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sonrce
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                      DEFINITION
                                                                                                                                                                                                                                                                                                                                                                                         BI981053/c
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ACCESSION
VERSION
KEYWORDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AUTHORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMMENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SOURCE
                                                                                Score:
                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                  Locus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Email: thisse@igpmc.ustrasbg.fr
Exarin: thisse@igpmc.ustrasbg.fr
Ext from a cDNA of a gene whose expression is spatially restricted during embrygonesis. We have esstablished its expression pattern during embrygonesis. We have esstablished its expression pattern on zebrafish embryos from the gastrula stage to 2 days of development. The corresponding data are available on the zebrafish community database at http://zfin.org/cDNA library preparation: B. Riggleman. DNA Sequencing by:IGBMC sequencing facility. Clone distribution: zebrafish international resource center at the University of Oregon (Institute of Neuroscience, 1254 University of Oregon. Eugene. OR 97403-1254).
Seq primer: T3 ATTAACCCTCACTAAAGGGA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CB417286 564 bp mRNA linear EST 27-MAR-2003 STR00763 gastrula stage cDNA library Danio rerio cDNA clone CB380 5' similar to CCAAT/enhancer binding protein beta, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygil; Neopterygil; Teleostel; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1, rue Laurent Fries, BP163, CU de Strasbourg, 67404 Illkirch Cedex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               l (bases 1 to 564)
Loppin, B., Steffan, T., Kempf, J., Heyer, V., Thisse, C. and Thisse, B.
Expression of the zebrafish genome during embryogenesis
Unpublished
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     /dD_Aic. ... /clone="CB380" /clone="CB380" /dov_stage="Gastrula stage embryos" /clone_lib="gastrula stage cDNA library" /note="Vector: Lambda Zap; Site_l: EcoRi; Site_2: XhoI; Oligo dT cDNA library constructed from RNA pooled from gastrula stage zebrafish embryos" | 155 q 128 t
- oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT/T3 vector. Library is non-normalized Library was constructed by Ning Wu. NOTE: This clone is available royalty-free through LLNL: contact the IMAGE Consortium (info.llnl.gov)for further information.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Contact: Thisse B
Institut de Genetique et de Biologie Moleculaire et Cellulaire
CNRS, INSERM, ULP
                                                                                                                                                                                                                                                                                                                Conservative:
Mismatches:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             355 GGGCAGTGGTTGTCATGGGAAGAGTTGTACTG 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GlySerValValIleValGlyArgIleValLeu 11
                                                                                                                                                                                                                                                                      Length:
Matches:
                                                                                                                                                                                                                                                                                                                                                                   Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Danio rerio"
/moi_type="mRNA"
/db_xref="taxon:7955"
                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-965-594-26 (1-11) x BE556909 (1-502)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CB417286.1 GI:29294527
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Danio rerio (zebrafish)
Danio rerio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Tel: 33 3 88 65 33 60
Fax: 33 3 88 65 32 01
                                                                                                                                                                                                                                                                                                                100.00%
63.64%
83.67%
10
                                                                                                                                                                                                                                                                      434
                                                                                                                                                                                                                                                                                                                                             Best Local Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CB417286
                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                Alignment Scores:
                                                                                                                                                                                                                                                                                                                                                              Query Match:
DB:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           source
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BASE COUNT
ORIGIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LOCUS
DEFINITION
                                                                                                                                                                                                                                                                           Pred. No.:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ORGANISM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AUTHORS
TITLE
JOURNAL
COMMENT
                                                                                                                                                                         SASE COUNT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VERSION
KEYWORDS
SOURCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ACCESSION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CB417286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ð
```

```
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
721: 314 286 1800
Fax: 314 286 1810
Fax: 314 286 1810
Fax: 314 286 1810
Faxi: 316 286 1810
F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       B1981053 584 bp mRNA linear EST 26-JUL-2002 fu38b04.x1 zebrafish adult brain Danio rerio cDNA clone IMAGE:5332062 3' similar to TR:097894 097894 CCAAT/ENHANCER BINDING PROTEIN BETA ;, mRNA sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cyprinidere; Eyprinides Danio Pyprinides Danio Pases I to 584;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 (bases 1 to 584)
Clark, M. Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S.L., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Rohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /dev_stage_"adult".

/lab_host="E. coli DH10B"
//lone_lib="zebrafish adult brain"
/clone_lib="zebrafish adult brain"
/clote="vector: pZIPLOX; Site_1: Not1; Site_2: Sal1;
Original library was constructed in lambdaZIPLOX. M
excision of the cDNA library was performed to yield
pZIPLOX plasmids. Insert check was done in original
library.
a 151 118 g 156 t
564
0000
                                                                                                             Conservative:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mismatches:
Indels:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 GlySerValValIleValGlyArglleValLeu 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /sex="mixed male and female"
/tissue_type="brain"
                                                   Matches:
          Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Seg primer: -400P
High quality sequence stop: 422.
Location/Qualiflers
1. 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:5332062"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WashU Zebrafish EST Project 1998
                                                                                                                                                                                                                                                                                                                                                 US-09-965-594-26 (1-11) x CB417286 (1-564)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             BI981053
BI981053.1 GI:16369292
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Danio rerio (zebrafish)
Danio rerio
                                              41.00
100.008
63.648
83.678
```

Alignment Scores:		
Pred. No.:	531	Length:
Score:	41.00	Matches:
Percent Similarity:	100.008	Conservative:
Best Local Similarity:	63.648	Mismatches:
Match:	873 678	Todolo.

9 2 2

DB: 12 Indels: 12 Gaps: US-09-965-594-26 (1-11) x BI981053 (1-584)

Search completed: August 31, 2003, 04:27:56 Job time: 112.667 secs

us-09-965-594-26.rst